PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12Q 1/68, C07K 14/65, A01K 67/02

A2

(11) International Publication Number: WO 00/36143

(43) International Publication Date: 22 June 2000 (22.06.00)

(21) International Application Number: PCT/EP99/10209

(22) International Filing Date: 16 December 1999 (16.12.99)

(30) Priority Data:
98204291.3 16 December 1998 (16.12.98) EP

(71) Applicants (*for all designated States except US*): UNIVERSITY OF LIEGE [BE/BE]; 20 Bd de Colonster, B-4000 Liege (BE). MELICA HB [SE/SE]; Andersson, Leif, Bergagatan 30, S-752 39 Uppsala (SE). SEGHERSGENTEC N.V. [BE/BE]; Kapelbaan 15, B-9255 Buggenhout (BE).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): ANDERSSON, Leif [SE/SE]; Bergagatan 30, S-752 39 Uppsala (SE). GEORGES, Michel [BE/BE]; Rue Vieux Tige 24, B-3161 Villers-aux-Tours (BE). SPINCEMAILLE, Geert [BE/BE]; Sint Denijsstraat 26, B-8550 Zwevegem (BE).

(74) Agent: OTTEVANGERS, S., U.; Vereenigde, Nieuwe Parklaan 97, NL-2587 BN The Hague (NL).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: SELECTING ANIMALS FOR PARENTALLY IMPRINTED TRAITS

(57) Abstract

The invention relates to methods to select breeding animals or animals destined for slaughter for having desired genotypic or potential phenotypic properties, in particular related to muscle mass and/or fat deposition. The invention provides a method for selecting a pig for having desired genotypic or potential phenotypic properties comprising testing a sample from said pig for the presence of a quantitative trait locus (QTL) located at a Sus scrofa chromosome 2 mapping at position 2p1.7.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

Title: Selecting animals for parentally imprinted traits.

The invention relates to methods to select breeding animals or animals destined for slaughter for having desired genotypic or potential phenotypic properties, in particular related to muscle mass and/or fat deposition. Breeding schemes for domestic animals have so far focused on farm performance traits and carcass quality. This has resulted in substantial improvements in traits like reproductive success, milk production, lean/fat ratio, prolificacy, growth rate and feed efficiency. Relatively simple performance test data have been the basis for these improvements, and selected traits were assumed to be influenced by a large number of genes, each of small effect (the infinitesimal gene model). There are now some important changes occurring in this area. First, the breeding goal of some breeding organisations has begun to include meat quality attributes in addition to the "traditional" production traits. Secondly, evidence is accumulating that current and new breeding goal traits may involve relatively large effects (known as major genes), as opposed to the infinitesimal model that has been relied on so far.

Modern DNA-technologies provide the opportunity to exploit these major genes, and this approach is a very promising route for the improvement of meat quality, especially since direct meat quality assessment is not viable for potential breeding animals. Also for other traits such as lean/fat ratio, growth rate and feed efficiency, modern DNA technology can be very effective. Also these traits are not always easy to measure in the living animal.

The evidence for several of the major genes originally obtained using segregation analysis, i.e. without any DNA marker information. Afterwards molecular studies were performed to detect the location of these

genes on the genetic map. In practice, and except for alleles of very large effect, DNA studies are required to dissect the genetic nature of most traits of economic importance. DNA markers can be used to localise genes or alleles responsible for qualitative traits like coat colour, and they can also be used to detect genes or alleles with substantial effects on quantitative traits like growth rate, IMF etc. In this case the approach is referred to as QTL (quantitative trait locus) mapping, wherein a QTL comprises at least a part of the nucleic acid genome of an animal where genetic information capable of influencing said quantitative trait (in said animal or in its offspring) is located. Information at DNA level can not only help to fix a specific major gene in a population, but also assist in the selection of a quantitative trait which is already selected for. Molecular information in addition to phenotypic data can increase the accuracy of selection and therefore the selection response.

Improving meat quality or carcass quality is not just about changing levels of traits like tenderness or marbling, but it is also about increasing uniformity. The existence of major genes provides excellent opportunities for improving meat quality because it allows large steps to be made in the desired direction. Secondly, it will help to reduce variation, since we can fix relevant genes in our products. Another aspect is that selecting for major genes allows differentiation for specific markets. Studies are underway in several species, particularly, pigs, sheep, deer and beef cattle.

In particular, intense selection for meat production has resulted in animals with extreme muscularity and leanness in several livestock species. In recent years it has become feasible to map and clone several of the genes causing these phenotypes, paving the way towards more efficient marker assisted selection, targeted drug development (performance enhancing products) and transgenesis. Mutations in the ryanodine receptor (Fuji

et al, 1991; MacLennan and Phillips, 1993) and myostatin (Grobet et al, 1997; Kambadur et al, 1997; McPherron and Lee, 1997) have been shown to cause muscular hypertrophies in pigs and cattle respectively, while genes with major effects on muscularity and/or fat deposition have for instance been mapped to pig chromosome 4 (Andersson et al, 1994) and sheep chromosome 18 (Cocket et al, 1996).

However, although there have been successes in identifying QTLs, the information is currently of limited use within commercial breeding programmes. Many workers in this field conclude that it is necessary to identify the particular genes underlying the QTL. This is a substantial task, as the QTL region is usually relatively large and may contain many genes. Identification of the relevant genes from the many that may be involved thus remains a significant hurdle in farm animals.

The invention provides a method for selecting a domestic animal for having desired genotypic or potential phenotypic properties comprising testing said animal for the presence of a parentally imprinted qualitative or quantitative trait locus (QTL). Herein, a domestic animal is defined as an animal being selected or having been derived from an animal having been selected for having desired genotypic or potential phenotypic properties.

Domestic animals provide a rich resource of genetic and phenotypic variation, traditionally domestication involves selecting an animal or its offspring for having desired genotypic or potential phenotypic properties. This selection process has in the past century been facilitated by growing understanding and utilisation of the laws of Mendelian inheritance. One of the major problems in breeding programs of domestic animals is the negative genetic correlation between reproductive capacity and production traits. This is for example the case in cattle (a high milk production generally results

in slim cows and bulls) poultry, broiler lines have a low level of egg production and layers have generally very low muscle growth), pigs (very prolific sows are in general fat and have comparatively less meat) or sheep (high prolific breeds have low carcass quality and vice versa). The invention now provides that knowledge of the parental imprinting character of various traits allows to select for example sire lines homozygous for a paternally imprinted QTL for example linked with muscle production or growth; the selection for such traits can thus be less stringent in dam lines in favour of the reproductive quality. The phenomenon of genetic or parental imprinting has never been utilised in selecting domestic animals, it was never considered feasible to employ this elusive genetic characteristic in practical breeding programmes. The invention provides a breeding programme, wherein knowledge of the parental imprinting character of a desired trait, as demonstrated herein, results in a breeding programme, for example in a BLUP programme, with a modified animal model. This increases the accuracy of the breeding value estimation and speeds up selection compared to conventional breeding programmes. Until now, the effect of a parentally imprinted trait in the estimation of a conventional BLUP programme was neglected; using and understanding the parental character of the desired trait, as provided by the invention, allows selecting on parental imprinting, even without DNA testing. For example, selecting genes characterised by paternal imprinting is provided to help increase uniformity; a (terminal) parent homozygous for the "good or wanted" alleles will pass them to all offspring, regardless of the other parent's alleles, and the offspring will all express the desired parent's alleles. This results in more uniform offspring. Alleles that are interesting or favourable from the maternal side or often the ones that have opposite effects to alleles from the paternal side. For example, in meat animals such as pigs alleles linked with meat quality traits such as inta-

muscular fat or muscle mass could be fixed in the dam lines while alleles linked with reduced back fat could be fixed in the sire lines. Other desirable combinations are for example fertility and/or milk yield in the female line with growth rates and/or muscle mass in the male lines.

In a preferred embodiment, the invention provides a method for selecting a domestic animal for having desired genotypic or potential phenotypic properties comprising testing a nucleic acid sample from said animal for the presence of a parentally imprinted quantitative trait locus (QTL). A nucleic acid sample can in general be obtained from various parts of the animal's body by methods known in the art. Traditional samples for the purpose of nucleic acid testing are blood samples or skin or mucosal surface samples, but samples from other tissues can be used as well, in particular sperm samples, oocyte or embryo samples can be used. In such a sample, the presence and/or sequence of a specific nucleic acid, be it DNA or RNA, can be determined with methods known in the art, such as hybridisation or nucleic acid amplification or sequencing techniques known in the art. The invention provides testing such a sample for the presence of nucleic acid wherein a QTL or allele associated therewith is associated with the phenomenon of parental imprinting, for example where it is determined whether a paternal or maternal allele of said QTL is capable of being predominantly expressed in said animal.

The purpose of breeding programs in livestock is to enhance the performances of animals by improving their genetic composition. In essence this improvement accrues by increasing the frequency of the most favourable alleles for the genes influencing the performance characteristics of interest. These genes are referred to as QTL. Until the beginning of the nineties, genetic improvement was achieved via the use of biometrical methods, but without molecular knowledge of the underlying QTL.

Since the beginning of the nineties and due to recent developments in genomics, it is conceivable to identify the QTL underlying a trait of interest. The invention now provides identifying and using parentally imprinted QTLs which are useful for selecting animals by mapping quantitative trait loci. Again, the phenomenon of genetic or paternal imprinting has never been utilised in selecting domestic animals, it was never considered feasible to employ this elusive genetic characteristic in practical breeding programmes. For example Kovacs and Kloting (Biochem. Mol. Biol. Int. 44:399-405, 1998), where parental imprinting is not mentioned, and not suggested, found linkage of a trait in female rats, but not in males, suggesting a possible sex specificity associated with a chromosomal region, which of course excludes parental imprinting, a phenomenon wherein the imprinted trait of one parent is preferably but gender-aspecifically expressed in his or her offspring.

The invention provides the initial localisation of a parentally imprinted QTL on the genome by linkage analysis with genetic markers, and the actual identification of the parentally imprinted gene(s) and causal mutations therein. Molecular knowledge of such a parentally imprinted QTL allows for more efficient breeding designs herewith provided. Applications of molecular knowledge of parentally imprinted QTLs in breeding programs include: marker assisted segregation analysis to identify the segregation of functionally distinct parentally imprinted QTL alleles in the populations of interest, marker assisted selection (MAS) performed within lines to enhance genetic response by increasing selection accuracy, selection intensity or by reducing the generation interval using the understanding of the phenomonon of parental imprinting, marker assisted introgression (MAI) to efficiently transfer favourable parentally imprinted QTL alleles from a donor to a recipient population, genetic engineering of the identified parentally QTL and genetic modification of the breeding stock using transgenic technology, development

of performance enhancing products using targeted drug development exploiting molecular knowledge of said QTL.

The inventors undertook two independent experiments to determine the practical use of parental imprinting of a QTL.

In a first experiment, performed in a previously described Piétrain x Large White intercross, the likelihood of the data were computed under a model of paternal (paternal allele only expressed) and maternal imprinting (maternal allele only expressed) and compared with the likelihood of the data under a model of a conventional "Mendelian" QTL. The results strikingly demonstrated that the QTL was indeed paternally expressed, the QTL allele (Piétrain or Large White) inherited from the F1 sow having no effect whatsoever on the carcass quality and quantity of the F_2 offspring. It was seen that very significant lodscores were obtained when testing for the presence of a paternally expressed QTL, while there was no evidence at all for the segregation of a QTL when studying the chromosomes transmitted by the sows. The same tendency was observed for all traits showing that the same imprinted gene is responsible for the effects observed on the different traits. Table 1 reports the maximum likelihood (ML) phenotypic means for the F_2 offspring sorted by inherited paternal QTL allele.

In a second experiment performed in the Wild Boar X Large White intercross, QTL analyses of body composition, fatness, meat quality, and growth traits was carried out with the chromosome 2 map using a statistical model testing for the presence of an imprinting effect. Clear evidence for a paternally expressed QTL located at the very distal tip of 2p was obtained (Fig. 2; Table1). The clear paternal expression of a QTL is illustrated by the least squares means which fall into two classes following the population origin of the paternally inherited allele (Table 1). For a given paternally imprinted QTL, implementation of marker assisted segregation analysis, selection (MAS) and introgression (MAI), can be performed

using genetic markers that are linked to the QTL, genetic markers that are in linkage disequilibrium with the QTL, or using the actual causal mutations within the QTL.

Understanding the parent-of-origin effect characterising a QTL allows for its optimal use in breeding programs. Indeed, marker assisted segregation analysis under a model of parental imprinting will yield better estimates of QTL allele effects. Moreover it allows for the application of specific breeding schemes to optimally exploit a QTL. In one embodiment of the invention, the most favourable QTL alleles would be fixed in breeding animal lines and for example used to generate commercial, crossbred males by marker assisted selection (MAS, within lines) and marker assisted introgression (MAI, between lines). In another embodiment, the worst QTL alleles would be fixed in the animal lines used to generate commercial crossbred females by MAS (within lines) and MAI (between lines).

In a preferred embodiment of the invention, said animal is a pig. Note for example that the invention provides the insight that today half of the offspring from commercially popular Piétrain$_x$ Large White crossbred boars inherit an unfavourable Large White muscle mass QTL as provided by the invention causing considerable loss, and the invention now for example provides the possibility to select the better half of the population in that respect. However, it is also possible to select commercial sow lines enriched with the in the boars unfavourable alleles, allowing to equip the sows with other alleles more desirable for for example reproductive purposes.

In a preferred embodiment of a method provided by the invention, said QTL is located at a position corresponding to a QTL located at chromosome 2 in the pig. For example, it is known form comparative mapping data between pig and human, including bidirectional chromosome painting, that SSC2p is homologous to HSA11pter-q13[11,12]. HSA11pter-q13 is known to harbour a

cluster of imprinted genes: IGF2, INS2, H19, MAH2, $P57^{KIP2}$, K_vLQTL1, Tapal,/CD81, Orctl2, Imptl and Ipl. The cluster of imprinted genes located in HSA11pter-q13 is characterised by 8 maternally expressed genes H19, MASH2, $P57^{KIP2}$, K_vLQTL1, TAPA1/CD81, ORCTL2, IMPT1 and IP1, and two paternally expressed genes: IGF2 and INS. However, Johanson et al (Genomics 25:682-690, 1995) and Reik et al (Trends in Genetics, 13:330-334, 1997) show that the whereabouts of these loci in various animals are not clear. For example, the HSA11 and MMU7 loci do not correspond among each other, the MMU7 and the SSC2 loci do not correspond, whereas the HSA11 and SSC2 loci seem to correspond, and no guidance is given where one or more of for example the above identified parentally expressed individual genes are localised on the three species' chromosomes.

Other domestic animals, such as cattle, sheep, poultry and fish, having similar regions in their genome harbouring such a cluster of imprinted genes or QTLs, the invention herewith provides use of these orthologous regions of other domestic animals in applying the phenomenon of parental imprinting in breeding programmes. In pigs, said cluster is mapped at around position 2p1.7 of chromosome 2, however, a method as provided by the invention employing (fragments of) said maternally or paternally expressed orthologous or homologous genes or QTLs are advantageously used in other animals as well for breeding and selecting purposes. For example, a method is provided wherein said QTL is related to the potential muscle mass and/or fat deposition, preferably with limited effects on other traits such as meat quality and daily gain of said animal or wherein said QTL comprises at least a part of an insulin-like growth factor-2 (IGF2) allele. Reik et al (Trends in Genetics, 13:330-334, 1997) explain that this gene in humans is related to Beckwith-Wiedemann syndrome, an apparently parentally imprinted disease syndrome most commonly seen with human foetuses, where the gene has an important role in prenatal

development. No relationship is shown or suggested with postnatal development relating to muscle development or fatness in (domestic) animals.

In a preferred embodiment, the invention provides a method for selecting a pig for having desired genotypic or potential phenotypic properties comprising testing a sample from said pig for the presence of a quantitative trait locus (QTL) located at a Sus scrofa chromosome 2 mapping at position 2p1.7. In particular, the invention relates to the use of genetic markers for the telomeric end of pig chromosome 2p in marker selection (MAS) of a parentally imprinted Quantitative Trait Locus (QTL) affecting carcass yield and quality in pigs. Furthermore, the invention relates to the use of genetic markers associated with the IGF2 locus in MAS in pigs, such as polymorphisms and microsatelites and other characterising nucleic acid sequences shown herein, such as shown in figures 4 to 10. In a preferred embodiment, the invention provides a QTL located at the distal tip of *Sus scrofa* chromosomes 2 with effects on varies measurements of carcass quality and quantity, particularly muscle mass and fat deposition.

In a first experiment, a QTL mapping analysis was performed in a Wild Boar X Large White intercross counting 200 F_2 individuals. The F_2 animals were sacrificed at a live eight of at least 80 kg or at a maximum age of 190 days. Phenotypic data on birth weight, growth, fat deposition, body composition, weight of internal organs, and meat quality were collected; a detailed description of the phenotypic traits are provided by Andersson *et al*[1] and Andersson-Eklund *et al*[4].

A QTL (without any significant effect on back-fat thickness) at an unspecified locus on the proximal end of chromosome 2 with moderate effect on muscle mass, and located about 30cM away from the parentally imprinted QTL reported here, was previously reported by the inventors; whereas the QTL as now provided has a very large effect, explaining at least 20-30% of variance, making the QTL of

the present invention commercially very attractive, which is even more so because the present QTL is parentally imprinted. The marker map of chromosome 2p was improved as part of this invention by adding microsatellite markers in order to cover the entire chromosome arm. The following microsatellite markers were used: *Swc9*, *Sw2443*, *Sw2623*, and *Swr2516*, all from the distal end of 2p[7]. QTL analyses of body composition, fatness, meat quality, and growth traits were carried out with the new chromosome 2 map. Clear evidence for a QTL located at the very distal tip of 2p was obtained (Fig. 1; Table 1). The QTL had very large effects on lean meat content in ham and explained an astonishing 30% of the residual phenotypic variance in the F_2 population. Large effects on the area of the longissumus dorsi muscle, on the weight of the heart, and on back-fat thickness (subcutaneous fat) were also noted. A moderate effect on one meat quality trait, reflectance value, was indicated. The QTL had no significant effect on abdominal fat, birth weight, growth, weight of liver, kidney, or spleen (data not shown). The Large White allele at this QTL was associated with larger muscle mass and reduced back-fat thickness consistent with the difference between this breed and the Wild Boar population.

In a second experiment, QTL mapping was performed in a Piétrain X Large White intercross comprising 1125 F_2 offspring. The Large White and Piétrain parental breeds differ for a number of economically important phenotypes. Piétrains are famous for their exceptional muscularity and leanness [10](Figure 2, while Large Whites show superior growth performance. Twenty-one distinct phenotypes measuring growth performance (5), muscularity (6), fat deposition (6), and meat quality (4), were recorded on all F_2 offspring. In order to map QTL underlying the genetic differences between these breeds, the inventors undertook a whole genome scan using microsatellite markers on an initial sample of 677 F_2 individuals. The following microsatellite marker map was used to analyse

chromosome 2;:SW2443, SWC9 and SW2623, SWR2516-(0,20)-SWR783-(0,29)-SW240-(0,20)-SW776-(0,08)-S0010-(0,04)-SW1695-(0,36)-SWR308. Analysis of pig chromosome 2 using a Maximum Likelihood multipoint algorithm, revealed highly significant lodscores (up to 20) for three of the six phenotypes measuring muscularity (% lean cuts, % ham, % loin) and three of the six phenotypes measuring fat deposition (back-fat thickness (BFT), % backfat, % fat cuts) at the distal end of the short arm of chromosome 2 (Figure 1). Positive lodscores were obtained in the corresponding chromosome region for the remaining six muscularity and fatness phenotypes, however, not reaching the experiment-wise significance threshold)(α=5%. There was no evidence for an effect of the corresponding QTL on growth performance (including birth weight) or recorded meat quality measurements (data not shown). To confirm this finding, the remaining sample of 355 F_2 offspring was genotyped for the four most distal 2p markers and QTL analysis performed for the traits yielding the highest lodscores in the first analysis. Lodscores ranged from 2.1 to 7.7, clearly confirming the presence of a major QTL in this region. Table 2 reports the corresponding ML estimates for the three genotypic means as well as the residual variance. Evidence based on marker assisted segregation analysis points towards residual segregation at this locus within the Piétrain population.

These experiments therefore clearly indicated the existence of a QTL with major effect on carcass quality and quantity on the telomeric end of pig chromosome arm 2p; the likely existence of an allelic series at this QTL with at least three alleles: Wild-Boar < Large White < Piétrain, and possibly more given the observed segregation within the Piétrain breed.

The effects of the identified QTL on muscle mass and fat deposition are truly major, being of the same magnitude of those reported for the *CRC* locus though apparently without the associated deleterious effects on meat quality. We estimate that both loci jointly explain

close to 50% of the Piétrain versus Large White breed difference for muscularity and leanness. The QTL had very large effects on lean meat content in ham and explained an astonishing 30% of the residual phenotypic variance in the F_2 population. Large effects on the area of the longissumus dorsi muscle, on the weight of the heart, and on back-fat thickness (subcutaneous fat) were also noted. A moderate effect on one meat quality trait, reflectance value, was indicated. The QTL had no significant effect on abdominal fat, birth weight, growth, weight of liver, kidney, or spleen (data not shown). The Large White allele at this QTL, when compared to the Wild Boar allele, was associated with larger muscle mass and reduced back-fat thickness consistent with the difference between this breed and the Wild Boar population. The strong imprinting effect observed for all affected traits shows that a single causative locus is involved. The pleiotropic effects on skeletal muscle mass and the size of the heart appear adaptive from a physiological point of view as a larger muscle mass requires a larger cardiac output.

In a further embodiment, the invention provides a method for selecting a pig for having desired genotypic or potential phenotypic properties comprising testing a sample from said pig for the presence of a quantitative trait locus (QTL) located at a Sus scrofa chromosome 2 mapping at position 2p1.7., wherein said QTL comprises at least a part of a Sus scrofa insulin-like growth factor-2 (IGF2) allele or a genonic area closely related thereto, such as polymorphisms and microsatelites and other characterising nucleic acid sequences shown herein, such as shown in figures 4 to 10. The important role of *IGF2* for prenatal development is well-documented from knock-out mice as well as from its causative role in the human Beckwith-Wiedemann syndrome. This invention demonstrates an important role for the *IGF2*-region also for postnatal development.

To show the role of Igf2 the inventors performed the following three experiments:

A genomic IGF2 clone was isolated by screening a porcine BAC library. FISH analysis with this BAC clone gave a strong consistent signal on the terminal part of chromosome 2p.

A polymorphic microsatellite is located in the 3'UTR of IGF2 in mice (GenBank U71085), humans (GenBank S62623), and horse (GenBank AF020598). The possible presence of a corresponding porcine microsatellite was investigated by direct sequencing of the IFG2 3'UTR using the BAC clone. A complex microsatellite was identified about 800bp downstream of the stop codon; a sequence comparison revealed that this microsatellite was identical to a previously described anonymous microsatellite, *Swc9*[6]. This marker was used in the initial QTL mapping experiments and its location on the genetic map correspond with the most likely position of the QTL both in the Piétrain X Large White and in the Large White x Wild Boar pedigree.

Analysis of skeletal muscle and liver cDNA from 10-week old foetuses heterozygous for a nt241 (G-A) transversion in the second exon of the porcine IGFII gene and SWC9, shows that the IGFII gene is imprinted in these tissues in the pig as well and only expressed from the paternal allele.

Based on a published porcine adult liver cDNA sequence[16], the inventors designed primer pairs allowing to amplify the entire *IgfII* coding sequence with 222 bp of leader and 280 bp of trailor sequence from adult skeletal muscle cDNA. Piétrain and Large White RT-PCR products were sequenced indication that the coding sequences are identical in both breeds and with the published sequence. However, a G⌫A transition was found in the leader sequence corresponding to exon 2 in man. Following conventional nomenclature, this polymorphism will be referred to as *nt241*(G-A). We developed a screening test for this single nucleotide polymorphism

9(SNP) based on the ligation amplification reaction (LAR), allowing us to genotype our pedigree material. Based on these data, *IgfII* was shown to colocalize with the SWC9 microsatellite marker (θ=0%), therefore virtually coinciding with the most likely position of the QTL, and well within the 95% support interval for the QTL. Subsequent sequence analysis demonstrated that the microsatellite marker SWC9 is actually located within the 3'UTR of the *IgfII* gene.

As previously mentioned, the knowledge of this QTL provides a method for the selection of animals such as pigs with improved carcass merit. Different embodiments of the invention are envisaged, including: marker assisted segregation analysis to identify the segregation of functionally distinct QTL alleles in the populations of interest; marker assisted selection (MAS) performed within lines to enhance genetic response by increasing selection accuracy, selection intensity or by reducing the generation interval; marker assisted introgression (MAI) to efficiently transfer favourable QTL alleles from a donor to a recipient population, thereby enhancing genetic response in the recipient population. Implementation of embodiments marker assisted segregation analysis, selection (MAS) and introgression (MAI), can be performed using genetic markers that are linked to the QTL; genetic markers that are in linkage disequilibrium with the QTL, the actual causal mutations within the QTL.

In a further embodiment, the invention provides a method for selecting a pig for having desired genotypic or potential phenotypic properties comprising testing a sample from said pig for the presence of a quantitative trait locus (QTL) located at a Sus scrofa chromosome 2 mapping at position 2p1.7., wherein said QTL is paternally expressed, i.e. is expressed from the paternal allele. In man and mouse, *Igf2* is known to be imprinted and to be expressed exclusively from the paternal allele in several tissues· Analysis of skeletal muscle cDNA from

pigs heterozygous for the SNP and/or SWC9, shows that the same imprinting holds in the pig as well. Understanding the parent-of-origin effect characterising the QTL as provided by the invention now allows for its optimal use in breeding programs. Indeed, today half of the offspring from commercially popular Piétrain x Large White crossbred boars inherit the unfavourable Large White allele causing considerable loss. Using a method as provide by the invention avoids this problem.

The invention furthermore provides an isolated and/or recombinant nucleic acid or functional fragment derived thereof comprising a parentally imprinted quantitative trait locus (QTL) or fragment thereof capable of being predominantly expressed by one parental allele. Having such a nucleic acid as provided by the invention available allows constructing transgenic animals wherein favourable genes are capable of being exclusively or predominantly expressed by one parental allele, thereby equipping the offspring of said animal homozygous for a desired trait with desired properties related to that parental allele that is expressed.

In a preferred embodiment, the invention provides an isolated and/or recombinant nucleic acid or fragment derived thereof comprising a synthetic parentally imprinted quantitative trait locus (QTL) or functional fragment thereof derived from at least one chromosome. Synthetic herein describes a parentally expressed QTL wherein various elements are combined that originate from distinct locations from the genome of one or more animals. The invention provides recombinant nucleic acid wherein sequences related to parental imprinting of one QTL are combined with sequences relating to genes or favourable alleles of a second QTL. Such a gene construct is favourably used to obtain transgenic animals wherein the second QTL has been equipped with paternal imprinting, as opposed to the inheritance pattern in the native animal from which the second QTL is derived. Such a second QTL can for example be derived from the same

chromosome where the parental imprinting region is located, but can also be derived from a different chromosome from the same or even a different species. In the pig, such a second QTL can for example be related to an oestrogen receptor (ESR)-gene (Rothschild et al, PNAS, 93, 201-201, 1996) or a FAT-QTL (Andersson, Science, 263, 1771-1774, 1994) for example derived from an other pig chromosome, such as chromosome 4. A second or further QTL can also be derived from another (domestic) animal or a human.

The invention furthermore provides an isolated and/or recombinant nucleic acid or functional fragment derived thereof at least partly corresponding to a QTL of a pig located at a Sus scrofa chromosome 2 mapping at position 2p1.7 wherein said QTL is related to the potential muscle mass and/or fat deposition of said pig and/or wherein said QTL comprises at least a part of a Sus scrofa insulin-like growth factor-2 (IGF2) allele, preferably at least spanning a region between INS and H19, or preferably derived from a domestic pig, such as a Pietrain, Meishan, Duroc, Landrace or Large White, or from a Wild Boar. For example, a genomic *IGF2* clone was isolated by screening a porcine BAC library. FISH analysis with this BAC clone gave a strong consistent signal on the terminal part of chromosome 2p. A polymorphic microsatellite is located in the 3'UTR of *IGF2* in mice (GenBank U71085), humans (GenBank S62623), and horse (GenBank AF020598). The possible presence of a corresponding porcine microsatellite was investigated by direct sequencing of the *IGF2* 3'UTR using the BAC clone. A complex microsatellite was identified about 800 bp downstream of the stop codon; a sequence comparison revealed that this microsatellite is identical to a previously described anonymous microsatellite, *Swc9*. PCR primers were designed and the microsatellite (*IGF2ms*) was found to be highly polymorphic with three different alleles among the two Wild Boar founders and another two

among the eight Large White founders. *IGF2ms* was fully informative in the intercross as the breed of origin as well as the parent of origin could be determined with confidence for each allele in each F_2 animal.

A linkage analysis using the intercross pedigree was carried out with *IGF2ms* and the microsatellites *Sw2443*, *Sw2623*, and *Swr2516*, all from the distal end of 2p[7]. *IGF2* was firmly assigned to 2p by highly significant lod scores (e.g. $Z=89.0$, $\theta=0.003$ against *Swr2516*). Multipoint analyses, including previously typed chromosome 2 markers, revealed the following order of loci (sex-average map distances in Kosambi cM): *Sw2443/Swr2516*-0.3-*IGF2*-14.9-*Sw2623*-10.3-*Sw256*. No recombinant was observed between *Sw2443* and *Swr2516*, and the suggested proximal location of *IGF2* in relation to these loci is based on a single recombinant giving a lod score support of 0.8 for the reported order. The most distal marker in our previous QTL study, *Sw256*, is located about 25 cM from the distal end of the linkage group.

The invention furthermore provides use of a nucleic acid or functional fragment derived thereof according to the invention in a method according to the invention. In a preferred embodiment, use of a method according to invention is provided to select a breeding animal or animal destined for slaughter, or embryos or semen derived from these animals for having desired genotypic or potential phenotypic properties. In particular, the invention provides such use wherein said properties are related to muscle mass and/or fat deposition. The QTL as provided by the invention may be exploited or used to improve for example lean meat content or back-fat thickness by marker assisted selection within populations or by marker assisted introgression of favorable alleles from one population to another. Examples of marker assisted selection using the QTL as provided by the invention are use of marker assisted segregation analysis

with linked markers or with markers in disequilibrium to identify functionally distinct QTL alleles. Furthermore, identification of a causative mutation in the QTL is now possible, again leading to identify functionally distinct QTL alleles. Such functionally distinct QTL alleles located at the distal tip of chromosome 2p with large effects on skeletal muscle mass, the size of the heart, and on back-fat thickness are also provided by the invention. The observation of a similar QTL effect in a Large White x Wild Boar as well as in a Piétrain x Large White intercross provides proof of the existence of a series of at least three distinct functional alleles. Moreover, preliminary evidence based on marker assisted segregation analysis points towards residual segregation at this locus within the Piétrain population (data not shown). The occurrence of an allelic series as provided by the invention allows identifying causal polymorphisms which - based on the quantitative nature of the observed effect - are unlikely to be gross gene alterations but rather subtle regulatory mutations. The effects on muscle mass of the three alleles rank in the same order as the breeds in which they are found i.e. Piétrain pigs are more muscular than Large White pigs that in turn have higher lean meat content than Wild Boars. The invention furthermore provides use of the alleles as provided by the invention for within line selection or for marker assisted introgression using linked markers, markers in disequilibrium or alleles comprising causative mutations.

The invention furthermore provides an animal selected by using a method according to the invention. For example, a pig characterised in being homozygous for an allele in a QTL located at a Sus scrofa chromosome 2 mapping at position 2p1.7 can now be selected and is thus provided by the invention. Since said QTL is related to the potential muscle mass and/or fat deposition of said pig and/or said QTL comprises at least a part of a Sus scrofa insulin-like growth factor-2 (IGF2) allele, it is

possible to select promising pigs to be used for breeding or to be slaughtered. In particular an animal according to the invention which is a male is provided. Such a male, or its sperm or an embryo derived thereof can advantageously be used in breeding animals for creating breeding lines or for finally breeding animals destined for slaughter. In a preferred embodiment of such use as provided by the invention, a male, or its sperm, deliberately selected for being homozygous for an allele causing the extreme muscular hyperthrophy and leanness, is used to produce offspring heterozygous for such an allele. Due to said allele's paternal expression, said offspring will also show the favourable traits for example related to muscle mass, even if the parent female has a different genetic background. Moreover, it is now possible to positively select the female(s) for having different traits, for example related to fertility, without having a negative effect on the muscle mass trait that is inherited from the allele from the selected male. For example, earlier such males could occasionally be seen with Piétrain pigs but genetically it was not understood how to most profitably use these traits in breeding programmes.

Furthermore, the invention provides a transgenic animal, sperm and an embryo derived thereof, comprising a synthetic parentally imprinted QTL or functional fragment thereof as provided by the invention, i.e. it is provided by the invention to introduce a favourable recombinant allele; for example introduce the oestrogen receptor locus related to increased litter size of an animal homozygously in a parentally imprinted region of a grandparent animal (for example the father of a hybrid sow if the region was paternally imprinted and the grandparent was a boar); to introduce a favourable fat-related allele or muscle mass-related recombinant allele in a paternally imprinted region, and so on. Recombinant alleles that are interesting or favourable from the maternal side or often the ones that have opposite effects to alleles from the paternal side. For example,

in meat animals such as pigs recombinant alleles linked with meat quality traits such as intra-muscular fat or muscle mass could be fixed in the dam lines while recombinant alleles linked with reduced back fat could be fixed in the sire lines. Other desirable combinations are for example fertility and/or milk yield in the female line with growth rates and/or muscle mass in the male lines.

The invention is further explained in the detailed description without limiting the invention.

Detailed description.

Example 1: Wild Boar x Large White intercrosses

Methods

Isolation of an *IGF2* BAC clone and fluorescent *in situ* hybridization (FISH). *IGF2* primers (F:5'-GGCAAGTTCTTCCGCTAATGA-3' and R:5'-GCACCGCAGAATTACGACAA-3') for PCR amplification of a part of the last exon and 3'UTR were designed on the basis of a porcine IGF2 cDNA sequence (GenBank X56094). The primers were used to screen a porcine BAC library and the clone 253G10 was isolated. Crude BAC DNA was prepared as described[24]. The BAC DNA was linearized with *Eco*RV and purified with QIAEXII (QIAGEN GmbH, Germany). The clone was labeled with biotin-14-dATP using the GIBCO-BRL Bionick labeling system (BRL18246-015). Porcine metaphase chromosomes were obtained from pokeweed (Seromed) stimulated lymphocytes using standard techniques. The slides were aged for two days at room temperature and then kept at -20°C until use. FISH analysis was carried out as previously described[25]. The final concentration of the probe in the hybridization mix was 10 ng/µl. Repetitive sequences were suppressed with standard concentrations of porcine

genomic DNA. After post-hybridization washing, the biotinylated probe was detected with two layers of avidin-FITC (Vector A-2011). The chromosomes were counterstained with 0.3 mg/ml DAPI (4,6-Diamino-2-phenylindole; Sigma D9542), which produced a G-banding like pattern. No posthybridization banding was needed, since chromosome 2 is easily recognized without banding. A total of 20 metaphase spreads were examined under an Olympus BX-60 fluorescence microscope connected to an IMAC-CCD S30 video camera and equipped with an ISIS 1.65 (Metasystems) software.

Sequence, microsatellite, and linkage analysis.

About two µg of linearized and purified BAC DNA was used for direct sequencing with 20 pmoles of primers and BigDye Terminator chemistry (Perkin Elmer, USA). DNA sequencing was done from the 3′ end of the last exon towards the 3′ end of the UTR until a microsatellite was detected. A primer set (F:5′-GTTTCTCCTGTACCCACACGCATCCC-3′ and R:5′-Fluorescein- CTACAAGCTGGGCTCAGGG-3′) was designed for the amplification of the *IGF2* microsatellite which is about 250 bp long and located approximately 800 bp downstream from the stop codon. The microsatellite was PCR amplified using fluorescently labeled primers and the genotyping was carried out using an ABI377 sequencer and the GeneScan/Genotyper softwares (Perkin Elmer, USA). Two-point and multipoint linkage analysis were done with the Cri-Map software[26].

Animals and phenotypic data.

The intercross pedigree comprised two European Wild Boar males and eight Large White females, 4 F_1 males and 22 F_1 females, and 200 F_2 progeny[1]. The F_2 animals were sacrificed at a live weight of at least 80 kg or at a

maximum age of 190 days. Phenotypic data on birth weight, growth, fat deposition, body composition, weight of internal organs, and meat quality were collected; a detailed description of the phenotypic traits are provided by Andersson *et al.*[1] and Andersson-Eklund *et al.*[4]

Statistical analysis.

Interval mapping for the presence of QTL were carried out with a least squares method developed for the analysis of crosses between outbred lines[27]. The method is based on the assumption that the two divergent lines are fixed for alternative QTL alleles. There are four possible genotypes in the F_2 generation as regards the grandparental origin of the alleles at each locus. This makes it possible to fit three effects: additive, dominance, and imprinting[2]. The latter is estimated as the difference between the two types of heterozygotes, the one receiving the Wild Boar allele through an F_1 sire and the one receiving it from an F_1 dam. An F-ratio was calculated using this model (with 3 d.f.) versus a reduced model without a QTL effect for each cM of chromosome 2. The most likely position of a QTL was obtained as the location giving the highest F-ratio. Genome-wise significance thresholds were obtained empirically by a permutation test[28] as described[2]. The QTL model including an imprinting effect was compared with a model without imprinting (with 1 d.f.) to test whether the imprinting effect was significant.

The statistical models also included the fixed effects and covariates that were relevant for the respective traits; see Andersson-Eklund *et al.*[4] for a more detailed description of the statistical models used. Family was included to account for background genetic

effects and maternal effects. Carcass weight was included as a covariate to discern QTL effects on correlated traits, which means that all results concerning body composition were compared at equal weights. Least-squares means for each genotype class at the *IGF2* locus were estimated with a single point analysis using Procedure GLM of SAS[29]; the model included the same fixed effects and covariates as used in the interval mapping analyses. The QTL shows a clear parent of origin-specific expression and the map position coincides with that of the insulin-like growth factor II gene (*IGF2*), indicating *IGF2* as the causative gene. A highly significant segregation distortion (excess of Wild Boar-derived alleles) was also observed at this locus. The results demonstrate an important effect of the *IGF2* region on postnatal development and it is possible that the presence of a paternally expressed *IGF2*-linked QTL in humans and in rodent model organisms has so far been overlooked due to experimental design or statistical treatment of data. The study has also important implications for quantitative genetics theory and practical pig breeding.

IGF2 was identified as a positional candidate gene for this QTL due to the observed similarity between pig chromosome 2p and human chromosome 11p. A genomic *IGF2* clone was isolated by screening a porcine BAC library. FISH analysis with this BAC clone gave a strong consistent signal on the terminal part of chromosome 2p (Fig. 1). A polymorphic microsatellite is located in the 3'UTR of *IGF2* in mice (GenBank U71085), humans (GenBank S62623), and horse (GenBank AF020598). The possible presence of a corresponding porcine microsatellite was investigated by direct sequencing of the *IGF2* 3'UTR using the BAC clone. A complex microsatellite was identified about 800 bp downstream of the stop codon; a sequence comparison revealed that this microsatellite is identical

to a previously described anonymous microsatellite, *Swc9*[6]. PCR primers were designed and the microsatellite (*IGF2ms*) was found to be highly polymorphic with three different alleles among the two Wild Boar founders and another two among the eight Large White founders. *IGF2ms* was fully informative in the intercross as the breed of origin as well as the parent of origin could be determined with confidence for each allele in each F_2 animal.

A linkage analysis using the intercross pedigree was carried out with *IGF2ms* and the microsatellites *Sw2443*, *Sw2623*, and *Swr2516*, all from the distal end of 2p[7]. *IGF2* was firmly assigned to 2p by highly significant lod scores (e.g. $Z=89.0$, $\theta=0.003$ against *Swr2516*). Multipoint analyses, including previously typed chromosome 2 markers[8], revealed the following order of loci (sex-average map distances in Kosambi cM): *Sw2443/Swr2516*-0.3-*IGF2*-14.9-*Sw2623*-10.3-*Sw256*. No recombinant was observed between *Sw2443* and *Swr2516*, and the suggested proximal location of *IGF2* in relation to these loci is based on a single recombinant giving a lod score support of 0.8 for the reported order. The most distal marker in our previous QTL study, *Sw256*, is located about 25 cM from the distal end of the linkage group.

QTL analyses of body composition, fatness, meat quality, and growth traits were carried out with the new chromosome 2 map using a statistical model testing for the possible presence of an imprinting effect as expected for *IGF2*. Clear evidence for a paternally expressed QTL located at the very distal tip of 2p was obtained (Fig. 2; Table 1). The QTL had very large effects on lean meat content in ham and explained an astonishing 30% of the residual phenotypic variance in the F_2 population. Large effects on the area of the longissumus dorsi muscle, on the weight of the heart, and on back-fat thickness

(subcutaneous fat) were also noted. A moderate effect on one meat quality trait, reflectance value, was indicated. The QTL had no significant effect on abdominal fat, birth weight, growth, weight of liver, kidney, or spleen (data not shown). The Large White allele at this QTL was associated with larger muscle mass and reduced back-fat thickness consistent with the difference between this breed and the Wild Boar population. The strong imprinting effect observed for all affected traits strongly suggests a single causative locus. The pleiotropic effects on skeletal muscle mass and the size of the heart appear adaptive from a physiological point of view as a larger muscle mass requires a larger cardiac output. The clear paternal expression of this QTL is illustrated by the least squares means which fall into two classes following the population origin of the paternally inherited allele (Table 1). It is worth noticing though that there was a non-significant trend towards less extreme values for the two heterozygous classes, in particular for the estimated effect on the area of longissimus dorsi. This may be due to chance, but could have a biological explanation, e.g. that there is some expression of the maternally inherited allele or that there is a linked, non-imprinted QTL with minor effects on the traits in question.

The *IGF2*-linked QTL and the *FAT1* QTL on chromosome 4 [1, 9] are by far the two loci with the largest effect on body composition and fatness segregating in this Wild Boar intercross. The *IGF2* QTL controls primarily muscle mass whereas *FAT1* has major effects on fat deposition including abdominal fat, a trait that was not affected by the *IGF2* QTL (Fig. 2). No significant interaction between the two loci was indicated and they control a very large proportion of the residual phenotypic variance in the F_2 generation. A model including both QTLs explains 33.1% of the variance for percentage lean meat in ham, 31.3% for the percentage of lean meat plus bone in back, and 26.2%

for average back fat depth (compare with a model including only chromosome 2 effects, Table 1). The two QTLs must have played a major role in the response during selection for lean growth and muscle mass in the Large White domestic pig.

A highly significant segregation distortion was observed in the *IGF2* region (excess of Wild Boar-derived alleles) as shown in Table 1 ($\chi 2$=11.7, d.f.=2; P=0.003). The frequency of Wild Boar-derived *IGF2* alleles was 59% in contrast to the expected 50% and there was twice as many "Wild Boar" as "Large White" homozygotes. This deviation was observed with all three loci at the distal tip and is thus not due to typing errors. The effect was also observed with other loci but the degree of distortion decreased as a function of the distance to the distal tip of the chromosome. Blood samples for DNA preparation were collected at 12 weeks of age and we are convinced that the deviation from expected Mendelian ratios was present at birth as the number of animals lost prior to blood sampling was not sufficient to cause a deviation of this magnitude. No other of the more than 250 loci analyzed in this pedigree show such a marked segregation distortion (L. Andersson, unpublished). The segregation distortion did not show an imprinting effect, as the frequencies of the two reciprocal types of heterozygotes were identical (Table 1). This does not exclude the possibility that the QTL effects and the segregation distortion are controlled by the same locus. The segregation distortion maybe due to meiotic drive favoring the paternally expressed allele during gametogenesis, as the F_1 parents were all sired by Wild Boar males. Another possibility is that the segregation distortion may be due to codominant expression of the maternal and paternal allele in some tissues and/or during a critical period of embryo development. Biallelic *IGF2* expression has been reported to occur to some extent

during human development[10, 11] and interestingly a strong influence of the parental species background on *IGF2* expression was recently found in a cross between *Mus musculus* and *Mus spretus*[12]. It is also interesting that a VNTR polymorphism at the insulin gene, which is very closely linked to *IGF2*, is associated with size at birth in humans[13]. It is possible that the *IGF2*-linked QTL in pigs has a minor effect on birth weight but in our data it was far from significant (Fig. 2) and there was no indication of an imprinting effect.

This study is an advance in the general knowledge concerning the biological importance of the *IGF2* locus. The important role of *IGF2* for prenatal development is well-documented from knock-out mice[14] as well as from its causative role in the human Beckwith-Wiedemann syndrome[15]. This study demonstrates an important role for the *IGF2*-region also for postnatal development. It should be stressed that our intercross between outbred populations is particularly powerful to detect QTL with a parent of origin-specific effect on a multifactorial trait. This is because multiple alleles (or haplotypes) are segregating and we could deduce whether a heterozygous F_2 animal received the Wild Boar allele from the F_1 male or female. It is quite possible that the segregation of a paternally expressed *IGF2*-linked QTL affecting a trait like obesity has been overlooked in human studies or in intercrosses between inbred rodent populations because of experimental design or statistical treatment of data. An imprinting effect cannot be detected in an intercross between two inbred lines as only two alleles are segregating at each locus. Our result has therefore significant bearings on the future analysis of the association between genetic polymorphism in the *insulin-IGF2* region and Type I diabetes[16], obesity[17], and variation in birth weight[13] in humans, as

well as for the genetic dissection of complex traits using inbred rodent models. A major impetus for generating an intercross between the domestic pig and its wild ancestor was to explore the possibilities to map and identify major loci that have responded to selection. We have now showed that two single QTLs on chromosome 2 (this study) and 4[1, 2] explain as much as one third of the phenotypic variance for lean meat content in the F_2 generation. This is a gross deviation from the underlying assumption in the classical infinitesimal model in quantitative genetics theory namely that quantitative traits are controlled by an infinite number of loci each with an infinitesimal effect. If a large proportion of the genetic difference between two divergent populations (e.g. Wild Boar and Large White) is controlled by a few loci, one would assume that selection would quickly fix QTL alleles with large effects leading to a selection plateau. However, this is not the experience in animal breeding programs or selection experiments where good persistent long-term selection responses are generally obtained, provided that the effective population size is reasonably large[18]. A possible explanation for this paradox is that QTL alleles controlling a large proportion of genetic differences between two populations may be due to several consecutive mutations; this may be mutations in the same gene or at several closely linked genes affecting the same trait. It has been argued that new mutations contribute substantially to long-term selection responses[19], but the genomic distribution of such mutations are unknown.

The search for a single causative mutation is the paradigm as regards the analysis of genetic defects in mice and monogenic disorders in humans. We propose that this may not be the case for loci that have been under selection for a large number of generations in domestic animals, crops, or natural populations. This hypothesis

predicts the presence of multiple alleles at major QTL. It gains some support from our recent characterization of porcine coat color variation. We have found that both the alleles for dominant white color and for black-spotting differ from the corresponding wild-type alleles by at least two consecutive mutations with phenotypic effects at the *KIT* and *MC1R* loci, respectively[20, 21]. In this context it is highly interesting that in the accompanying example we have identified a third allele at the *IGF2*-linked QTL. The effects on muscle mass of the three alleles rank in the same order as the breeds in which they are found i.e. Piétrain pigs are more muscular than Large White pigs that in turn have higher lean meat content than Wild Boars.

There are good reasons to decide that *IGF2* is the causative gene for the now reported QTL. Firstly, there is a perfect agreement in map localization (Fig. 2). Secondly, it has been shown that *IGF2* is paternally expressed in mice, humans, and now in pigs, like the QTL. There are several other imprinted genes in the near vicinity of *IGF2* in mice and humans (*Mash2, INS2, H19, KVLQT1, TAPA1/CD81*, and *CDKN1C/p57*^{KIP2}) but only *IGF2* is paternally expressed in adult tissues[22]. We believe that this locus provides a unique opportunity for molecular characterization of a QTL. The clear paternal expression can be used to exclude genes that do not show this mode of inheritance. Moreover, the presence of an allelic series should facilitate the difficult distinction between causative mutations and linked neutral polymorphism. We have already shown that there is no difference in coding sequence between *IGF2* alleles from Piétrain and Large White pigs suggesting that the causative mutations occur in regulatory sequences. An obvious step is to sequence the entire *IGF2* gene and its multiple promoters from the three populations. The recent

report that a VNTR polymorphism in the promoter region of the insulin (*INS*) gene affects *IGF2* expression[23] suggests that the causative mutations may be at a considerable distance from the *IGF2* coding sequence.

The results have several important implications for the pig breeding industry. They show that genetic imprinting is not an esoteric academic question but need to be considered in practical breeding programs. The detection of three different alleles in Wild Boar, Large White, and Piétrain populations indicates that further alleles at the *IGF2*-linked QTL segregate within commercial populations. The paternal expression of the QTL facilitates its detection using large paternal half-sib families as the female contribution can be ignored. The QTL is exploited to improve lean meat content by marker assisted selection within populations or by marker assisted introgression of favorable alleles from one population to another.

Example 2: Piétrain x Large White intercrosses

Methods

Pedigree material: The pedigree material utilized to map QTL was selected from a previously described Piétrain x Large White F2 pedigree comprising > 1,800 individuals[6,7]. To assemble this F2 material, 27 Piétrain boars were mated to 20 Large White sows to generate an F1 generation comprising 456 individuals. 31 F1 boars were mated to unrelated 82 F1 sows from 1984 to 1989, yielding a total of 1862 F2 offspring. F1 boars were mated on average to 7 females, and F1 sows to an average of 2,7 males. Average offspring per boar were 60 and per sow 23.

Phenotypic information: (i) Data collection: A total of 21 distinct phenotypes were recorded in the F2 generation[6,7]. These included:

- five growth traits: birth weight (g), weaning weight (Kg), grower weight (Kg), finisher weight (Kg) and average daily gain (ADG; Kg/day; grower to finsher period);
- two body proportion measurements: carcass length (cm); and a conformation score (0 to 10 scale; ref.6);
- ten measurements of carcass composition obtained by dissection of the chilled carcasses 24 hours after slaughter. These include measurements of muscularity: % ham (weight hams/carcass weight), % loin (weight loin/carcass weight), % shoulder (weight shoulder/carcass weight), % lean cuts (% ham + %loin + % shoulder); and measurements of fatness: average back-fat thickness (BFT; cm), % backfat (weight backfat/carcass weight), % belly (weight belly/carcass weight), % leaf fat (weight leaf fat/carcass weight), % jowl (weight jowl/carcass weight), and "% fat cuts" (% backfat + % belly + % leaft fat + % jowl).
- four meat quality measurements: pH $_{LD1}$ (*Longissimus dorsi* 1

hour after slaughter), pH $_{LD24}$ (*Longissimus dorsi* 24 hours after slaughter), pH $_{G1}$ (*Gracilis* 1 hour after slaughter) and pH $_{G24}$ (*Gracilis* 24 hours after slaughter). *(ii) Data processing:* Individual phenotypes were preadjusted for fixed effects (sire, dam, CRC genotype, sex, year-season, parity) and covariates (litter size, birth weight, weaning weight, grower weight, finisher weight) that proved to significantly affect the corresponding trait. Variables included in the model were selected by stepwise regression.

Marker genotyping: Primer pairs utilized for PCR amplification of microsatellite markers are as described[19]. Marker genotyping was performed as previously described[20]. Genotypes at the *CRC* and *MyoD* loci were determined using conventional methods as described[1,12]. The LAR test for the Igf2 SNP was developed according to Baron et al.[21] using a primer pair for PCR amplification (5'-CCCCTGAACTTGAGGACGAGCAGCC-3';5'-ATCGCTGTGGGCTGGGTGGGCTGCC-3') and a set of three primers for the LAR step (5'-FAM-CGCCCCAGCTGCCCCCCAG-3'; 5'-HEX-CGCCCCAGCTGCCCCCCAA-3'; 5'-CCTGAGCTGCAGCAGGCCAG-3').

Map construction: Marker maps were constructed using the TWOPOINT, BUILD and CHROMPIC options of the CRIMAP package[22]. To allow utilisation of this package, full-sib families related via the boar or sow were disconnected and treated independently. By doing so, some potentially usable information was neglected, yielding, however, unbiased estimates of recombination rates.

QTL mapping: (i) Mapping Mendelian QTL: Conventional QTL mapping was performed using a multipoint maximum likelihood method. The applied model assumed one segregating QTL per

chromosome, and fixation of alternate QTL alleles in the respective parental lines, Piétrain (P) and Large White (LW). A specific analysis program had to be developed to account for the missing genotypes of the parental generation, resulting in the fact that the parental origin of the F1 chromosomes could not be determined. Using a typical "interval mapping" strategy, an hypothetical QTL was moved along the marker map using user-defined steps. At each position, the likelihood *(L)* of the pedigree data was computed as:

$$L = \sum_{\varphi=1}^{2^r} \prod_{i=1}^{n} \sum_{G=1}^{4} (P(G|M_i,\theta,\varphi)P(y_i|G))$$

P or right chromosme *P*), there is a total of 2^r combinations for *r* F1 parents.

$\prod_{i=1}^{n}$ *n* F2

$\sum_{G=1}^{4}$ *i*th F2 offspring, over the four possible QTL genotypes: *P/P, P/LW, LW/P* and *LW/LW*

$P(G|M_i,0,\varphi)M_i$: the marker genotype of the *i*th F2 offspring and its F1 parents, (ii) : the vector of recombination rates between adjacent markers and between the hypothetical QTL and its flanking markers, and (iii)θ the considered marker-QTL phase combination of the F1 parents.

Recombination rates and marker linkage phase of F1 parents are assumed to be known when computing this probability. Both were determined using CRIMAP in the map construction phase (see above).

$P(y_i|G)\,y_i$) of offspring *i*, given the QTL genotype under consideration. This probability is computed from the normal density function:

$$P(y_i|G) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(y_i-\mu_G)^2}{2\sigma^2}}$$

$_G$ is the phenotypic mean of the considered QTL genotype (PP, PL, LP or LL) and σ^2 the residual variance σ^2 was considered to be the same for the four QTL genotypic classes.

The values of μ_{PP}, $\mu_{PL}=\mu_{LP}$, μ_{LL} and σ^2 maximizing L were determined using the GEMINI optimisation routine[23].

The likelihood obtained under this alternative H_1 hypothesis was compared with the likelihood obtained under the null hypothesis H_0 of no QTL, in which the phenotypic means of the four QTL genotypic classes were forced to be identical. The difference between the logarithms of the corresponding likelihoods yields a lodscore measuring the evidence in favour of a QTL at the corresponding map position.

*(ii) Significance thresholds:*Following Lander & Botstein[24], lodscore thresholds *(T)* associated with a chosen genome-wise significance level, were computed such that:

$$\alpha = (C + 9.21GT)\chi_2^2(4.6T)$$

C corresponds to the number of chromosomes (= 19), *G* corresponds to the length of the genome in Morgans (= 29), and χ_2^2 *(4.6T)* denotes one minus the cumulative distribution function of the chi-squared distribution with 2 d.f. Single point 2ln(LR) were assumed to be distributed as a chi-squared distribution with two degrees of freedom, as we were fitting both an additive and dominance component. To account for the fact that we were analysing multiple traits, significance levels were adjusted by applying a Bonferoni correction corresponding to the effective number of independent traits that were analyzed. This effective number was estimated at 16 following the approach described by Spelman et al.[25].

Altogether, this allowed us to set the lodscore threshold associated with an experiment-wise significance level of 5%

at 5.8. When attempting to confirm the identified QTL in an independent sample, the same approach was used, however, setting C at 1, G at 25cM and correcting for the analysis of 4.5 independent traits (as only six traits were analyzed in this sample). This yielded a lodscore threshold associated with a Type I error of 5% of 2.

(iii). Testing for an imprinted QTL: To test for an imprinted QTL, we assumed that only the QTL alleles transmitted by the parent of a given sex would have an effect on phenotype, the QTL alleles transmitted by the other parent being "neutral". The likelihood of the pedigree data under this hypothesis was computed using equation 1. To compute $P(y_i \mid G)$, however, the phenotypic means of the four QTL genotypes were set at $\mu_{PP} = \mu_{PL} = \mu_P$ and $\mu_{LP} = \mu_{LL} = \mu_L$ to test for a QTL for which the paternal allele only is expressed, and $\mu_{PP} = \mu_{LP} = \mu_P$ and $\mu_{PL} = \mu_{LL} = \mu_L$ to test for a QTL for which the maternal allele only is expressed. It is assumed in this notation that the first subscript refers to the paternal allele, the second subscript to the maternal allele. H_0 was defined as the null-hypothesis of no QTL, H_1 testing the presence of a Mendelian QTL; H_2 testing the presence of a paternally expressed QTL, and H_3 testing the presence of a maternally expressed QTL.

RT-PCR: Total RNA was extracted from skeletal muscle according to Chirgwin et al.[26]. RT-PCR was performed using the Gene-Amp RNA PCR Kit (Perkin-Elmer) The PCR products were purified using QiaQuick PCR Purification kit (Qiagen) and sequenced using Dye terminator Cycle Sequencing Ready Reaction (Perkin Elmer) and an ABI373 automatic sequencer.

In example 2 we report the identification of a QTL with major effect on muscle mass and fat deposition mapping to porcine 2p1.7 The QTL shows clear evidence for parental imprinting strongly suggesting the involvement of the *Igf2* locus.

A Piétrain X Large White intercross comprising 1125 F_2 offspring was generated as described[6,7]. The Large White and Piétrain parental breeds differ for a number of economically important phenotypes. Piétrains are famed for their exceptional muscularity and leannes [8] (Figure 2), while Large Whites show superior growth performance. Twenty-one distinct phenotypes measuring (i) growth performance (5), (ii) muscularity (6), (iii) fat deposition (6), and (iv) meat quality (4), were recorded on all F_2 offspring.

In order to map QTL underlying the genetic differences between these breeds, we undertook a whole genome scan using microsatellite markers on an initial sample of 677 F_2 individuals. Analysis of pig chromosome 2 using a ML multipoint algorithm, revealed highly significant lodscores (up to 20) for six of the 12 phenotypes measuring muscularity and fat deposition at the distal end of the short arm of chromosome 2 (Figure 3a). Positive lodscores were obtained for the remaining six phenotypes, however, not reaching the genome-wise significance threshold (= 5%). To confirm this finding, the remaining sample of 355 F_2 offspring was genotyped for the five most distal 2p markers and QTL analysis performed for the traits yielding the highest lodscores in the first analysis. Lodscores ranged from 2.1 to 7.7, clearly confirming the presence of a major QTL in this region. Table 2 reports the corresponding ML estimates for the three genotypic means as well as the corresponding residual variance.

Bidirectional chromosome painting establishes a correspondence between SSC2p and HSA11pter-q13[9,10]. At least

two serious candidate genes map to this region in man: the myogenic basic helix-loop-helix factor, *MyoD*, maps to HSA11p15.4, while *Igf2* maps to HSA11p15.5 *MyoD* is a well known key regulator of myogenesis and is one of the first myogenic markers to be switched on during development[11]. A previously described amplified sequence polymorphism in the porcine *MyoD* gene[12] proved to segregate in our F_2 material, which was entirely genotyped for this marker. Linkage analysis positioned the *MyoD* gene in the SW240-SW776 (odds > 1000) interval, therefore well outside the lod-2 drop off support interval for the QTL (figure 1). *Igf2* is known to enhance both proliferation and differentiation of myoblasts *in vitro*[13] and to cause a muscular hypertrophy when overexpressed *in vivo*. Based on a published porcine adult liver cDNA sequence[14], we designed primer pairs allowing us to amplify the entire *Igf2* coding sequence with 222 bp of leader and 280 bp of trailor sequence from adult skeletal muscle cDNA. Piétrain and Large White RT-PCR products were sequenced indicating that the coding sequences was identical in both breeds and with the published sequence. However, a G A transition was found in the leader sequence corresponding to exon 2 in man (Figure 4). We developed a screening test for this single nucleotide polymorphism (SNP) based on the ligation amplification reaction (LAR), allowing us to genotype our pedigree material. Based on these data, *Igf2* was shown to colocalize with the SWC9 microsatellite marker (= 0%), therefore located at approximately 2 centimorgan from the most likely position of the QTL and well within the 95% support interval for the QTL (figure 1). Subsequent sequence analysis demonstrated that the microsatellite marker SWC9 is actually located within the 3' UTR of the *Igf2* gene. Combined with available comparative mapping data for the PGA and FSH loci, these results suggest the occurrence of an interstitial

inversion of a chromosome segment containing *MyoD*, but not *Igf2* which has remained telomeric in both species.

Igf2 therefore appeared as a strong positional allele having the observed QTL effect. In man and mouse, *Igf2* is known to be imprinted and to be expressed exclusively from the paternal allele in several tissues[15]. Analysis of skeletal muscle cDNA from pigs heterozygous for the SNP and/or SWC9, shows that the same imprinting holds in this tissue in the pig as well (Figure 4). Therefore if *Igf2* were responsible for the observed effect, and knowing that only the paternal *Igf2* allele is expressed, one can predict that (i) the paternal allele transmitted by F1 boars (P or LW) would have an effect on phenotype of F2 offspring, (ii) the maternal allele transmitted by F1 sows (P or LW) would have no effect on phenotype of F2 offspring, and (iii) the likelihood of the data would be superior under a model of a bimodal (1:1) F2 population sorted by inherited paternal allele when compared to a conventional "Mendelian" model of a trimodal (1:2:1) F2 population. The QTL mapping programs were adapted in order to allow testing of the corresponding hypotheses. H_0 was defined as the null-hypothesis of no QTL, H_1 as testing for the presence of a Mendelian QTL, H_2 as testing for the presence of a paternally expressed QTL, and H_3 as testing for the presence of a maternally expressed QTL.

Figure 3 summarizes the obtained results. Figure 3a, 3b and 3c respectively show the lodscore curves corresponding to $\log_{10}(H_2/H_0)$, $\log_{10}(H_3/H_0)$ and $\log_{10}(H_2/H_1)$. It can be seen that very significant lodscores are obtained when testing for the presence of a paternally expressed QTL, while there is no evidence at all for the segregation of a QTL when studying the chromosomes transmitted by the sows. Also, the hypothesis of a paternally expressed QTL is significantly more likely ($\log_{10}(H_2/H_1) > 3$) than the hypothesis of a "Mendelian" QTL

for all examined traits. The fact that the same tendency is observed for all traits indicates that it is likely the same imprinted gene that is responsible for the effects observed on the different traits. Table 2 reports the ML phenotypic means for the F2 offspring sorted by inherited paternal QTL allele. Note that when performing the analysis under a model of a mendelian QTL, the Piétrain and Large White QTL alleles appeared to behave in an additive fashion, the heterozygous genotype exhibiting a phenotypic mean corresponding exactly to the midpoint between the two homzygous genotypes. This is exactly what one would predict when dealing with an imprinted QTL as halve of the heterozygous offspring are expected to have inherited the P allele from their sire, the other halve the LW allele.

These data therefore confirmed our hypothesis of the involvement of an imprinted gene expressed exclusively from the paternal allele. The fact that the identified chromosomal segment coincides precisely with an imprinted domain documented in man and mice strongly implicates the orthologous region in pigs. At least seven imprinted genes mapping to this domain have been documented (*Igf2, Ins2, H19, Mash2,* $p57^{KIP2}$, K_vLQTL1 and *TDAG51*) (ref. 15 and Andrew Feinberg, personal communication). Amongst these, only *Igf2* and *Ins2* are paternally expressed. While we cannot exclude that the observed QTL effect is due to an as of yet unidentified imprinted gene in this region, its reported effects on myogenesis *in vitro* and *in vivo*[13] strongly implicate *Igf2*. Particularly the muscular hypertrophy observed in transgenic mice overexpressing *Igf2* from a muscle specific promotor are in support of this hypothesis (Nadia Rosenthal, personal communication. Note that allelic variants of the *INS* VNTR have recently been shown to be associated

with size at birth in man[16], and that the same VNTR has been shown to affect the level of *Igf2* expression[17].

The observation of the same QTL effect in a Large White x Wild Boar intercross indicates the existence of a series of at least three distinct functional alleles. Moreover, preliminary evidence based on marker assisted segregation analysis points towards residual segregation at this locus within the Piétrain population (data not shown). The occurrence of an allelic series might be invaluable in identifying the causal polymorphisms which - based on the quantitatve nature of the observed effect - are unlikely to be gross gene alterations but rather subtle regulatory mutations.

The effects of the identified QTL on muscle mass and fat deposition are truly major, being of the same magnitude of those reported for the *CRC* locus[6,7] though apparently without the associated deleterious effects on meat quality. We estimate that both loci jointly explain close to 50% of the Piétrain versus Large White breed difference for muscularity and leanness. Understanding the parent-of-origin effect characterizing this locus will allow for its optimal use in breeding programs. Indeed, today half of the offspring from commercially popular Piétrain x Large White crossbred boars inherit the unfavourable Large White allele causing considerable loss.

The QTL described in this work is the second example of a gene affecting muscle development in livestock species that exhibits a non-mendelian inheritance pattern. Indeed, we have previously shown that the callipyge locus (related to the qualitative trait wherein muscles are doubled) is characterized by polar overdominance in which only the heterozygous individuals that inherit the CLPG mutation from their sire express the double-muscling phenotype[5]. This

demonstrates that parent-of-origin effects affecting genes underlying production traits in livestock might be relatively common.

Example 3:

Generating a reference sequence of IGF2 and flanking loci in the pig.

The invention provides an imprinted QTL with major effect on muscle mass mapping to the IGF2 locus in the pig, and use of the QTL as tool in marker assisted selection. To fine tune this tool for marker assisted selection, as well as to further identify a causal mutation, we have further generated a reference sequence encompassing the entire porcine IGF2 sequence as well as that from flanking genes.

To achieve this, we screened a porcine BAC library with IGF2 probes and identified two BACs. BAC-PIGF2-1 proved to contain the INS and IGF2 genes, while BAC-PIGF2-2 proved to contain the IGF2 and H19 genes. The NotI map as well as the relative position of the two BACs is shown in Figure 5. BAC-PIGF2-1 was shotgun sequenced using standard procedures and automatic sequencers. The resulting sequences were assembled using standard software yielding a total of 115 contigs. The corresponding sequences are reported in figure 6. Similarity searches were performed between the porcine contigs and the orthologous sequences in human. Significant homologies were detected for 18 contigs and are reported in Figure 7.

For BAC-PIGF2-2, the 24 Kb NotI fragment not present in BAC-PIGF2-1 was subcloned and sequenced using the EZ::TN transposon approach and ABI automatic sequencers. Resulting

sequences were assembled using the Phred-Phrap-Consed program suit, yielding seven distinct contigs (figure 8). The contig sequences were aligned with the corresponding orthologous human sequences using the compare and dotplot programs of the GCG suite. Figure 9 symmarizes the corresponding results.

Example 4: Identification of DNA sequence polymorphisms in the IGF2 and flanking loci.

Based on the reference sequence obtained as described in Example 1, we resequenced part of the IGF2 and flanking loci from genomic DNA isolated from Piétrain, Large White and Wild Boar individuals, allowing identification of DNA sequence polymorphisms such as reported in figure 10.

Legends to the figures

Fig. 1: Test statistic curves obtained in QTL analyses of chromosome 2 in a Wild Boar/Large White intercross. The graph plots the F ratio testing the hypothesis of a single QTL at a given position along the chromosome for the traits indicated. The marker map with the distances between markers in Kosambi centiMorgan is given on the X-axis. The horizontal lines represent genome-wise significant ($P<0.05$) and suggestive levels for the trait lean meat in ham; similar significance thresholds were obtained for the other traits.

Figure 2: Piétrain pig with characteristic muscular hypertrophy.

Figure 3: Lodscore curves obtained in a Piétrain x Large White intercross for six phenotypes measuring muscle mass and fat deposition on pig chromosome 2. The most likely positions of the *Igf2* and *MyoD* genes determined by linkage analysis with respect to the microsatellite marker map are shown. H_0 was defined as the null-hypothesis of no QTL, H_1 as testing for the presence of a Mendelian QTL, H_2 as testing for the presence of a paternally expressed QTL, and H_3 as testing for the presence of a maternally expressed QTL. 3a: $\log_{10}(H_1/H_0)$, 3b: $\log_{10}(H_2/H_0)$, 3c: $\log_{10}(H_3/H_0)$

Figure 4: A. Structure of the human *Igf2* gene according to ref. 17, with aligned porcine adult liver cDNA sequence as reported in ref. 16. The position of the *nt241(G-A)* transition and *Swc9* microsatellite are shown. B. The corresponding markers were used to demonstrate the monoallelic (paternal) expression of *Igf2* in skeletal muscle

and liver of 10-week old fetuses. PCR amplification of the *nt421(G-A)* polymorphism and *Swc9* microsatellite from genomic DNA clearly shows the heterozygosity of the fetus, while only the paternal allele is detected in liver cDNA (*nt421(G-A)* and *Swc9*) and muscle cDNA (*Swc9*). The absence of RT-PCR product for *nt421(G-A)* from in fetal muscle points towards the absence of mRNA including exon 2 in this tissue. Parental origin of the foetal alleles was determined from the genotypes of sire and dam (data not shown).

Figure 5: A NotI restriction map showing the relative position of BAC-PIGF2-1 (comprising INS and IGF2 genes), and BAC-PIGF2-2 (comprising IGF2 and H19 genes).

Figure 6: Nucleic acid sequences of contig 1 to contig 115 derived from BAC-PIGF2-1 which was shotgun sequenced using standard procedures and automatic sequencers.

Figure 7: Similarity between porcine contigs of figure 6 and orthologous sequences in human.

Figure 8 Nucleic acid sequences of contig 1 to contig 7 derived from BAC-PIGF2-2, (the 24 Kb NotI fragment not present in BAC-PIGF2-1) which was subcloned and sequenced using the EZ::TN transposon approach and ABI automatic sequencers.

Figure 9: Similarity between porcine contigs of figure 8 and orthologous sequences in human.

Figure 10: DNA sequence polymorphisms in the IGF2 and flanking loci from genomic DNA isolated from Piétrain, Large White and Wild Boar individuals.

REFERENCES

Literature cited with example 1

1. Andersson, L. *et al.* Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science* 263, 1771-1774 (1994).

2. Knott, S.A. *et al.* Multiple marker mapping of quantitative trait loci in a cross between outbred wild boar and Large White pigs. *Genetics* 149, 1069-1080 (1998).

3. Edfors-Lilja, I. *et al.* Mapping quantitative trait loci for immune capacity in the pig. *Journal of Immunology* 161, 829-835 (1998).

4. Andersson-Eklund, L. *et al.* Mapping quantitative trait loci for carcass and meat quality traits in a wild boar x Large White intercross. *Journal of Animal Science* 76, 694-700 (1998).

5. Fronicke, L., Chowdhary, B.P., Scherthan, H. & Gustavsson, I. A comparative map of the porcine and human genomes demonstrates ZOO-FISH and gene mapping-based chromosomal homologies. *Mamm Genome* 7, 285-90 (1996).

6. Alexander, L.J. *et al.* Physical assignments of 68 porcine cosmids and lambda clones containing microsatellites. *Mammalian Genome* 7, 368-372 (1996).

7. Rohrer, G.A. *et al.* A comprehensive map of the porcine genome. *Genome Research* 6, 371-391 (1996).

8. Marklund, L. *et al.* A comprehensive linkage map of the pig based on a wild pig-Large White intercross. *Anim Genet* 27, 255-69 (1996).

9. Marklund, L., Nyström, P.E., Stern, S., Andersson-Eklund, L. & Andersson, L. Quantitative trait loci for

fatness and growth on pig chromosome 4. *Heredity* In press(1998).

10. Ohlsson, R., Hedborg, F., Holmgren, L., Walsh, C. & Ekstrom, T.J. Overlapping patterns of IGF2 and H19 expression during human development: biallelic IGF2 expression correlates with a lack of H19 expression. *Development* 120, 361-368 (1994).

11. Ekström, T.J., Cui, H., Li, X. & Ohlsson, R. Promoter-specific IGF2 imprinting status and its plasticity during human liver development. *Development* 121, 309-316 (1995).

12. Hemberger, M. *et al.* H19 and Igf2 are expressed and differentially imprinted in neuroectoderm-derived cells in the mouse brain. *Dev. Genes Evol.* 208, 393-402 (1998).

13. Dunger, D.B. *et al.* Association of the INS VNTR with size at birth. *Nature Genetics* 19, 98-100 (1998).

14. DeChiara, T.M., Robertson, E.J. & Efstratiadis, A. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* 64, 849-859 (1991).

15. Sun, F.L., Dean, W.L., Kelsey, G., Allen, N.D. & Reik, W. Transactivation of Igf2 in a mouse model of Beckwith-Wiedemann syndrome. *Nature* 389, 809-815 (1997).

16. Davies, J.L. *et al.* A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* 371, 130-136 (1994).

17. O'Dell, S.D. *et al.* ApaI polymorphism in insulin-like growth factor II (IGF2) gene and weight in middle-aged males. *International Journal of Obesity* 21, 822-825 (1997).

18. Falconer, D.S. & Mackay, T.F.C. *Introduction to Quantitative Genetics*, (Longman, England, 1996).

19. Hill, W.G. Rates of change in quantitative traits from fixation of new mutations. *Proc Natl Acad Sci U S A* 79, 142-145 (1982).

20. Marklund, S. *et al.* Molecular basis for the dominant white phenotype in the domestic pig. *Genome Research* 8, 826-833 (1998).

21. Kijas, J.M.H. *et al.* Melanocortin receptor 1 (*MC1R*) mutations and coat color in the pig. *Genetics* In press(1998).

22. Beechey, C.V. personal communication (1998).

23. Paquette, J., Giannoukakis, N., Polychronakos, C., Vafiadis, P. & Deal, C. The *INS* 5' variable number of tandem repeats is associated with *IGF2* expression in humans. *Journal of Biological Chemistry* 273, 14158-14164 (1998).

24. Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular cloning : A laboratory manual.*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989).

25. Chowdhary, B.P., de la Sena, C., Harbitz, I., Eriksson, L. & Gustavsson, I. FISH on metaphase and interphase chromosomes demonstrates the physical order of the genes for GPI, CRC, and LIPE in pigs. *Cytogenetics Cell Genetics* 71, 175-178 (1995).

26. Green, P., Falls, K. & Crook, S. *Documentation for CRI-MAP, version 2.4.*, (Washington University School of Medicine, St Louise, MO, 1990).

27. Haley, C.S., Knott, S.A. & Elsen, J.M. Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics* 136, 1195-1207 (1994).

28. Churchill, G.A. & Doerge, R.W. Empirical threshold values for quantitative trait mapping. *Genetics* 138, 963-971 (1994).

29. Anonymous. *SAS version 6.10*, (SAS Institute Inc., Cary, NC., 1990).

References used with example 2:

1. Fuji, J.; Otsu, K.; Zorzato, F.; Deleon, S.; Khanna, V.K.; Weiler, J.E. O'Brien, P.J.; MacLennan, D.H. (1991). Identification of a mutation in the porcine ryanodine receptor associated with malignant hyperthermia. *Science* 253: 448-451.

2. MacLennan, D.H. & Phillips, M.S. (1993). Malignant hyperthermia. *Science* 256:789-794.

3. Grobet, L.; Royo Martin, L.J.; Poncelet, D.; Pirottin, D.; Brouwers, B.; Riquet, J.; Schoeberlein, A.; Dunner, S.; Ménissier, F.; Massabanda, J.; Fries, R.; Hanset, R.; Georges, M. (1997) A deletion in the myostatin gene causes double-muscling in cattle. *Nature Genetics* 17:71-74.

4. Andersson, L.; Haley, C.S.; Ellegren, H.; Knott, S.A.; Johansson, M.; Andersson, K.; Andersson-Eklund, L.; Edfors-Lilja, I.; Fredholm, M.; Hansson, I.; Håkansson, J.; Lundström, K. (1994). Genetic mapping of quantitative trait loci for growth and fatness in pigs. Science 263:1771-1774.

5. Cockett, N.; Jackson, S.; Shaw, T.; Farnir, F.; Berghmans, S.;Snowder, G.; Nielsen, D.; Georges, M. (1996). Polar overdominance at the ovine callipyge locus. *Science* 273:236-238

6. Hanset, R.; Dasnois, C.; Scalais, S.; Michaux, C.; Grobet, L. (1995). Genetypes at the locus for halothane sensitivity and performance in a Piétrain x Large White F2. Genet. Sel. Evol. 27: 63-76.

7. Hanset, R.; Dasnois, C.; Scalais, S.; Michaux, C.; Grobet, L. (1995). Introgression into the Piétrain genome of the normal allele at the locus for halothane sensitivity. Genet. Sel. Evol. 27: 77-88.

8. Olivier, L.; Lauvergne, J.J. (1967). A study of the inheritance of the muscular hypertrophy of the Piétrain pig: preliminary results. Annales de Médecine Vétérinaire 111: 104-109.

9. Rettenberger, G.; Klett, C.; Zechner, U.; Kunz, J.; Vogel, W.; Hameister, H. (1995). Visualisation of the conservation of synteny between humans and pigs by heterologous chromosome painting. *Genomics* 26: 372-378.

10. Goureau, A.; Yerle, M.; Schmitz, A.; Riquet, J.; Milan, D.; Pinton, P.; Frelat, G.; Gellin, J. (1996). Human and porcine correspondence of chromosome segments using bidirectional chromosome painting. *Genomics* 36:252-262.

11. Yun, K.; Wold, B. (1996). Skeletal muscle determination and differentiation: story of a core regulatory network and its context. Current Opinion in Cell Biology 8:877-889.

12. Knoll, A.; Nebola, M.; Dvorak, J.; Cepica, S. (1997). Detection of a DdeI PCR RFLP within intron 1 of the porcine MYOD1(MYF3) locus. *Animal Genetics* 28, 308-322.

13. Florini, J.R.; Ewton, D.Z.; McWade, F.J. (1995). IGFs, muscle growth, and myogenesis. Diabetes Review 3:73-92.

14. Catchpole, I.R.; Engström, W. (1990). Nucleotide sequence of a porcine insulin-like growth factor II cDNA. Nucleic Acids Research 18(21):6430.

15. Feil, R.; Moore, T.F.; Oswald, J.; Walter, J.; Sun, F.; Reik, W. (1997). The imprinted insulin like growth factor 2 gene. Pp70 In Genomic Imprinting. Eds. Reik & Surani. IRL Press at Oxford University Press.

16. Dunger, D.B.; Ong, K.K.L.; Huxtable, S.J.; Sherriff, A.; Woods, K.A.; Ahmed, M.L.; Golding, J.; Pembrey, M.E.; Ring, S.; the ALSPAC study team, Bennett, S.T.; Todd, J.A. (1998). Association of the INS VNTR with size at birth. *Nature Genetics* 19: 98-100.

17. Paquette J, Giannoukakis N, Polychronakos C, Vafiadis P, Deal C. (1998) The INS 5' variable number of tandem repeats is associated with IGF2 expression in humans. J. Biol Chem 273(23):14158-14164

18. Andersson-Eklund, L.; Marklund, L.; Lundström, K.; Haley, C.S.; Andersson, K.; Hansson, I.; Moller, M.; Andersson, L. (1998). Mapping Quantitative Trait Loci for carcass and meat quality traits in a Wild Boar x Large White intercross. *J. Anim. Sci.* 76:694-700.

19. Rohrer, G.A.; Alexander, L.J.; Hu, Z.; Keele, J.W.; Smith, T.P.; Beattie, C.W. (1996). A comprehensive map of the porcine genome. Genome Research, in the press.

20. Georges, M.; Nielsen, D.; Mackinnon, M.; Mishra, A.; Okimoto, R.; Pasquino, A.T.; Sargeant, L.S.; Sorensen, A.; Steele, M.R.; Zhao, X.; Womack, J.E.; Hoeschele, I. (1995). Mapping quantitative trait loci controlling milk production by exploiting progeny testing. *Genetics* 139: 907-920.

21. Baron, H.; Fung, S.; Aydin, A.; Bahring, S.; Luft, F.C.; Schuster, H. (1996). Oligonucleotide ligation assay (OLA) for the diagnosis of familial hypercholesterolemia. Nat. Biotechnol. 14(10):1279-1282.

22. Lander, E.; Green, P. (1987) Construction of multilocus genetic linkage maps in humans. Proceedings of National Academy of Science (USA) 84: 2363-2367.

23. Lalouel, J.M. (1983). Optimization of functions. Contrib. Epidemiol.Biostat. 4:235-259.

24. Lander, E.S. & Botstein, D. (1989). Mapping mendelian factots underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199.

25. Spelman RL, Coppieters W, Karim L, van Arendonk JAM, Bovenhuis H (1996) Quantitative trait loci analysis for five milk production traits on chromosome six in the dutch Holstein-Friesian population. Genetics 144:1799-1808.

26.Chirgwin, J.M.; Przybyla, A.E.; MacDonald, R.J.; Rutter, W.J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18:5294-5299

Table 1 Summary of QTL analysis for pig chromosome 2 in a Wild Boar/Large White intercross[1]

Trait	F ratio[2]		Map	Percent of F_2	Least squares means[5]			
	QTL	*Imprinting*	position[3]	variance[4]	W^P/W^M	W^P/L^M	L^P/W^M	L^P/L^M
					n=62	n=43	n=43	n=30
Body composition traits								
Lean meat in ham, %	24.4***	19.1***	0	30.6	63.6^a	64.2^a	66.4^b	67.3^b
Lean meat mass in ham, kg	18.1***	16.8***	1	24.3	4.69^a	4.72^a	4.94^b	5.02^b
Lean meat + bone in back, %	12.2**	9.6**	0	17.4	66.3^a	66.7^a	69.3^b	70.8^b
Longissimus muscle area, cm^2	10.3**	4.8*	1	15.4	31.9^a	33.0^a	34.5^b	35.2^b
Fatness traits								
Average back fat depth, mm	7.1*	8.7**	0	10.4	27.2^a	27.7^a	25.5^b	24.7^b
Weight of internal organs								
Heart, gram	9.7**	11.4***	0	14.4	226^a	225^a	238^b	244^b
Meat quality traits								
Reflectance value, EEL	5.7	6.1*	1	8.1	18.6^a	18.4^a	21.8^b	19.7^a

*P<0.05; **P<0.01; ***P<0.001

Table 1, continued

[1]Only the traits for which the QTL peak was in the *IGF2* region (0-10 cM) and the test statistic reached the nominal significance threshold of F=3.9 are included.

[2]"*QTL*" is the test statistic for the presence of a QTL under a genetic model with additive, dominance, and imprinting effects (3 d.f.) while "Imprinting" is the test statistic for the presence of an imprinting effect (1 d.f.), both obtained at the position of the QTL peak. Genome-wise significance thresholds, estimated by permutation, were used for the QTL test while nominal significance thresholds were used for the Imprinting test.

[3]In cM from the distal end of 2p; *IGF2* is located at 0.3 cM.

[4]The reduction in the residual variance of the F_2 population effected by inclusion of an imprinted QTL at the given position.

[5]Means and standard errors estimated at the *IGF2* locus by classifying the genotypes according to the population and parent of origin of each allele. *W* and *L* represent alleles derived from the Wild Boar and Large White founders, respectively; superscript *P* and *M* represent a paternal and maternal origin, respectively. Figures with different letters (superscript a or b) are significantly different at least at the 5% level, most of them are different at the 1% or 0.1% level.

Table 2 Maximum likelihood phenotypic means for the different F2 genotypes estimated under (i) a model of a mendelian QTL, and (ii) a model assuming an imprinted QTL.

Traits	Mendelian QTL				Imprinted QTL		
	$\mu_{LW/LW}$	$\mu_{LW/P}$	$\mu_{P/P}$	R	$\mu_{PAT/LW}$	$\mu_{PAT/P}$	R
BFT (cm)	2.98	2.84	2.64	0.27	2.94	2.70	0.27
% ham	21.10	21.56	22.15	0.83	21.23	21.95	0.83
% loin	24.96	25.53	26.46	0.91	25.12	26.14	0.93
% lean cuts	65.02	65.96	67.60	1.65	65.23	67.05	1.67
% backfat	6.56	6.02	5.33	0.85	6.43	5.56	0.85
% fat cuts	28.92	27.68	26.66	1.46	28.54	26.99	1.49

CLAIMS

1. A method for selecting a domestic animal for having desired genotypic properties comprising testing said animal for the presence of a parentally imprinted quantitative trait locus (QTL).

2. A method according to claim 1 further comprising testing a nucleic acid sample from said animal for the presence of a parentally imprinted quantitative trait locus (QTL).

3. A method according to claim 1 or 2 wherein in the pig said QTL is located at chromosome 2.

4. A method according to claim 2 or 3 wherein said QTL is mapping at around position 2p1.7.

5. A method according to claim 1 to 4 wherein said QTL is related to the potential muscle mass and/or fat deposition of said animal.

6. A method according to claim 5 wherein said QTL comprises at least a part of an insulin-like growth factor-2 (IGF2) gene.

7. A method according to anyone of claims 1 to 6 wherein in the pig said QTL comprises a marker characterised as nt241(G-A) or as Swc9, as identified in figure 4.

8. A method according to anyone of claims 1-7 wherein a paternal allele of said QTL is predominantly expressed in said animal.

9. A method according to anyone of claims 1-7 wherein a maternal allele of said QTL is predominantly expressed in said animal.

10. An isolated and/or recombinant nucleic acid comprising a parentally imprinted quantitative trait locus (QTL) or functional fragment derived thereof.

11. An isolated and/or recombinant nucleic acid comprising a synthetic parentally imprinted quantitative trait locus (QTL)

derived from at least one chromosome or functional fragment derived thereof.

12. A nucleic acid according to claim 10 or 11 at least partly derived from a Sus scrofa chromosome.

13. A nucleic acid according to claim 12 wherein said nucleic acid is at least partly derived from a Sus scrofa chromosome 2, preferably from a region mapping at around position 2p1.7.

14. A nucleic acid according to any one of claims 10 to 13 wherein said QTL is related to the potential muscle mass and/or fat deposition of said animal.

15. A nucleic acid according to any one of claims 10 to 14 wherein said QTL comprises at least a part of a insulin-like growth factor-2 (IGF2) gene.

16. A nucleic acid according to anyone of claims 10 to 15 wherein a paternal allele of said QTL is capable of being predominantly expressed.

17. A nucleic acid according to anyone of claims 10 to 16 wherein a maternal allele of said QTL is capable of being predominantly expressed.

18. Use of a nucleic acid or fragment derived thereof according to claim 10 in a method according to anyone of claims 1-9.

19. Use according to claim 18 to select a breeding animal or animal destined for slaughter for having desired genotypic or potential phenotypic properties.

20. Use according to claim 19 wherein said properties are related to muscle mass and/or fat deposition.

21. An animal such as pig selected by a use according to claim 18 to 20.

22. A animal according to claim 21 characterised in being homozygous for an allele at a paternally imprinted QTL, preferably located at a Sus scrofa chromosome 2 mapping at around position 2p1.7.

23. An animal according to claim 21 or 22 wherein said QTL is related to the potential muscle mass and/or fat deposition of

said pig and/or wherein said QTL comprises at least a part of a insulin-like growth factor-2 (IGF2) allele.

24. A transgenic animal comprising a nucleic acid according to anyone of claims 11 to 16.

25. An animal according to anyone of claims 21-24 which is a male.

26. Sperm or an embryo derived from an animal according to anyone of claims 21-25.

27. Use of a sperm or an embryo according to claim 26 in breeding animals destined for slaughter.

FIGURE 1

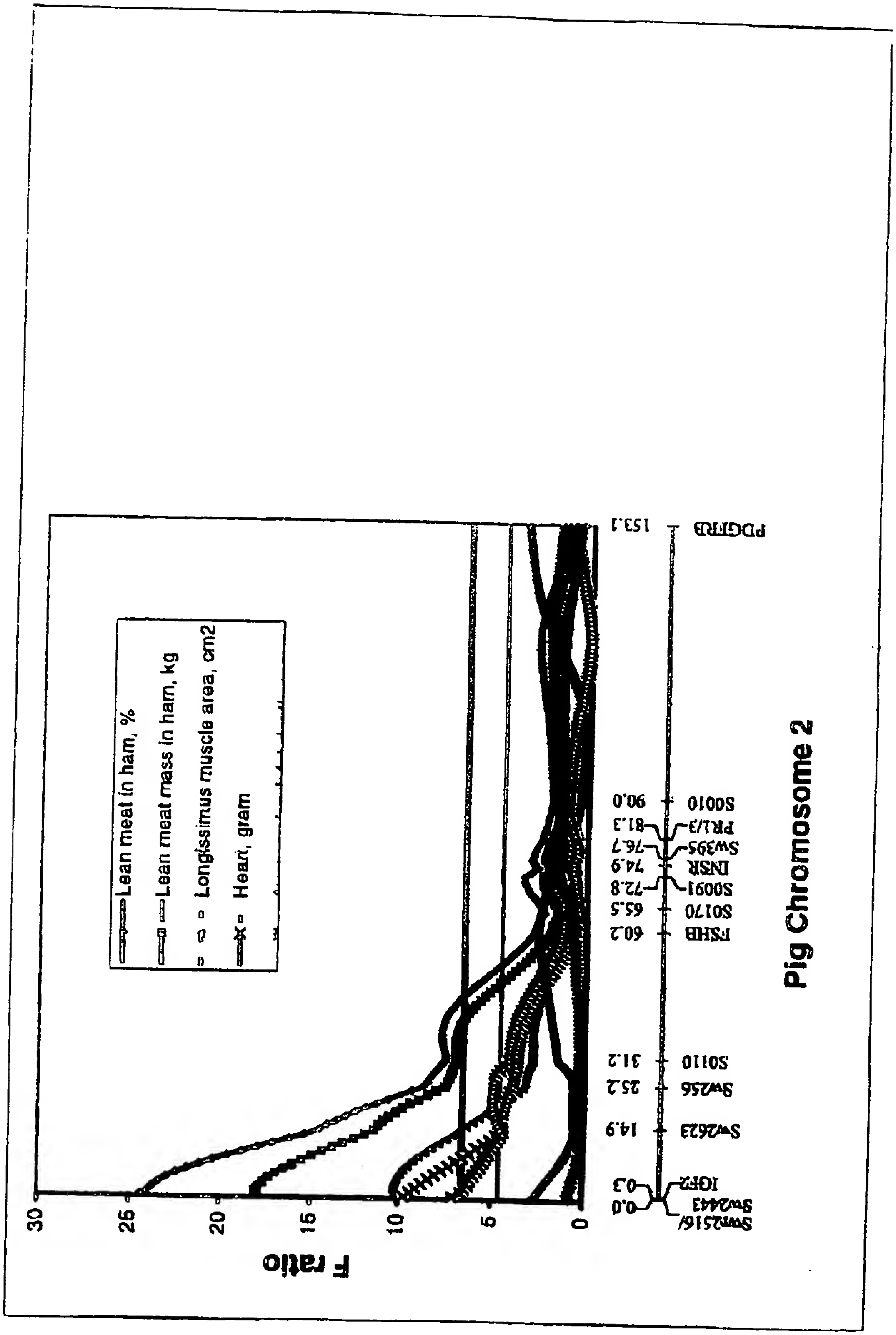
Lean meat in ham, %
Lean meat mass in ham, kg
Longissimus muscle area, cm2
Heart, gram
F ratio
Pig Chromosome 2
Sw2516/
Sw2443
IGF2
Sw2623
Sw256
S0110
FSHB
S0170
S0091
INSR
Sw395
PRLr3
S0010
PDGFRB

FIGURE 2

FIGURE 3A

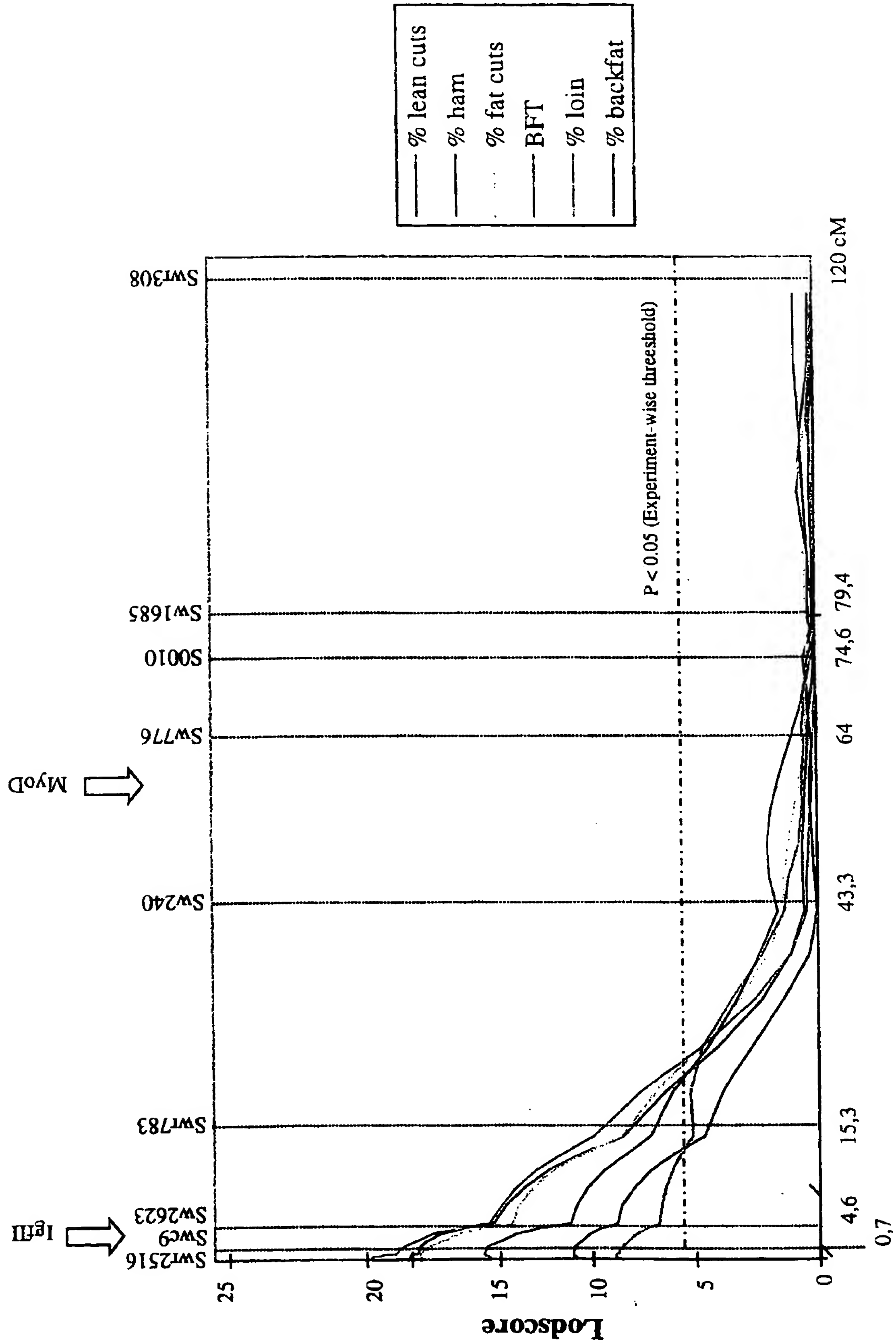
% lean cuts
% ham
% fat cuts
BFT
% loin
% backfat
Sw2516
Swc9
Sw2623
IgfII
Sw783
Sw240
MyoD
Sw776
S0010
Sw1685
Sw308
P < 0.05 (Experiment-wise threeshold)
0,7
4,6
15,3
43,3
64
74,6
79,4
120 cM
25
20
15
10
5
0
Lodscore

FIGURE 3B

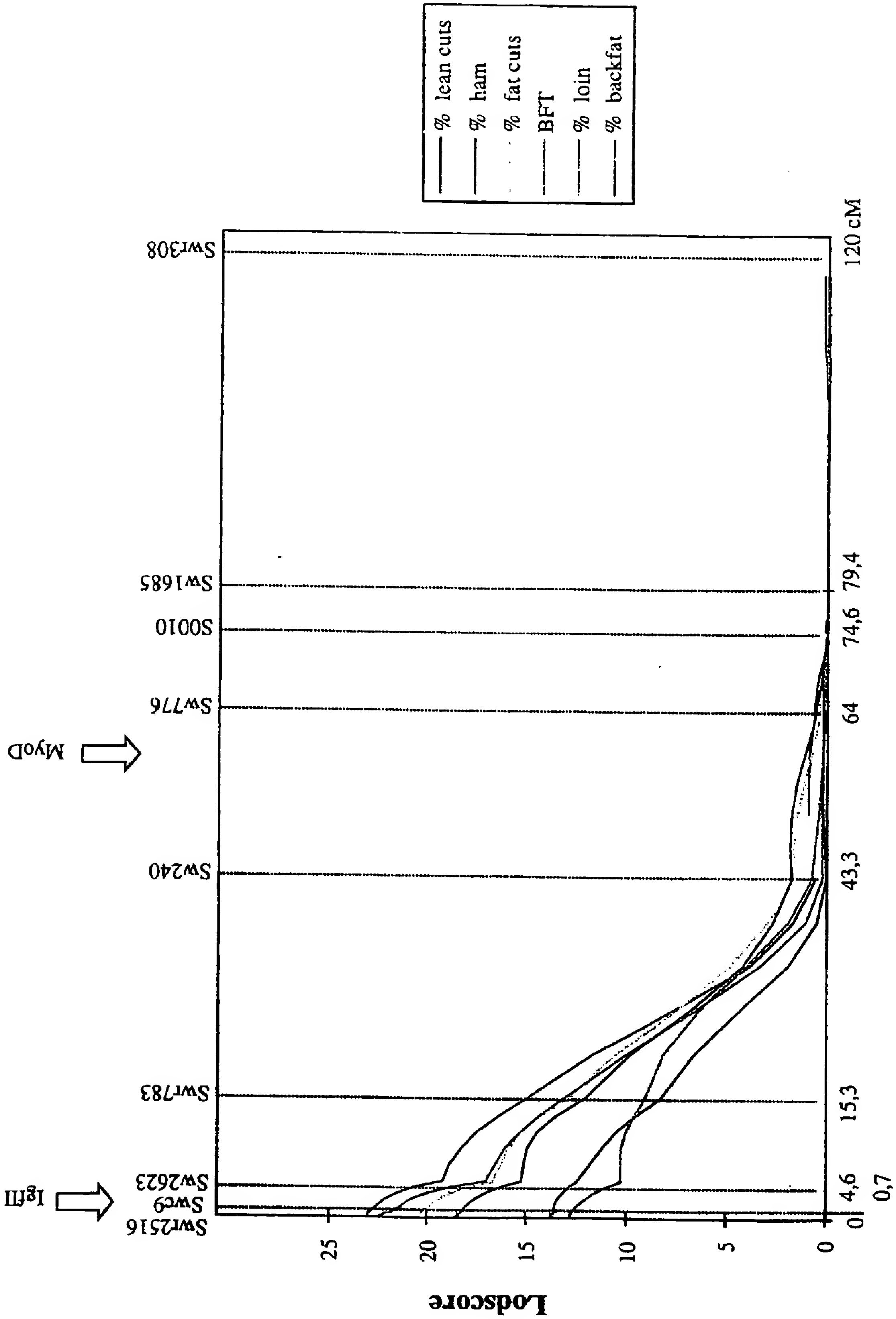
% lean cuts
% ham
% fat cuts
BFT
% loin
% backfat
Sw308
Sw1685
S0010
Sw776
MyoD
Sw240
Sw783
Sw2623
Sw9
Sw2516
IgfII
120 cM
79,4
74,6
64
43,3
15,3
4,6
0,7
0
25
20
15
10
5
0
Lodscore

FIGURE 3C

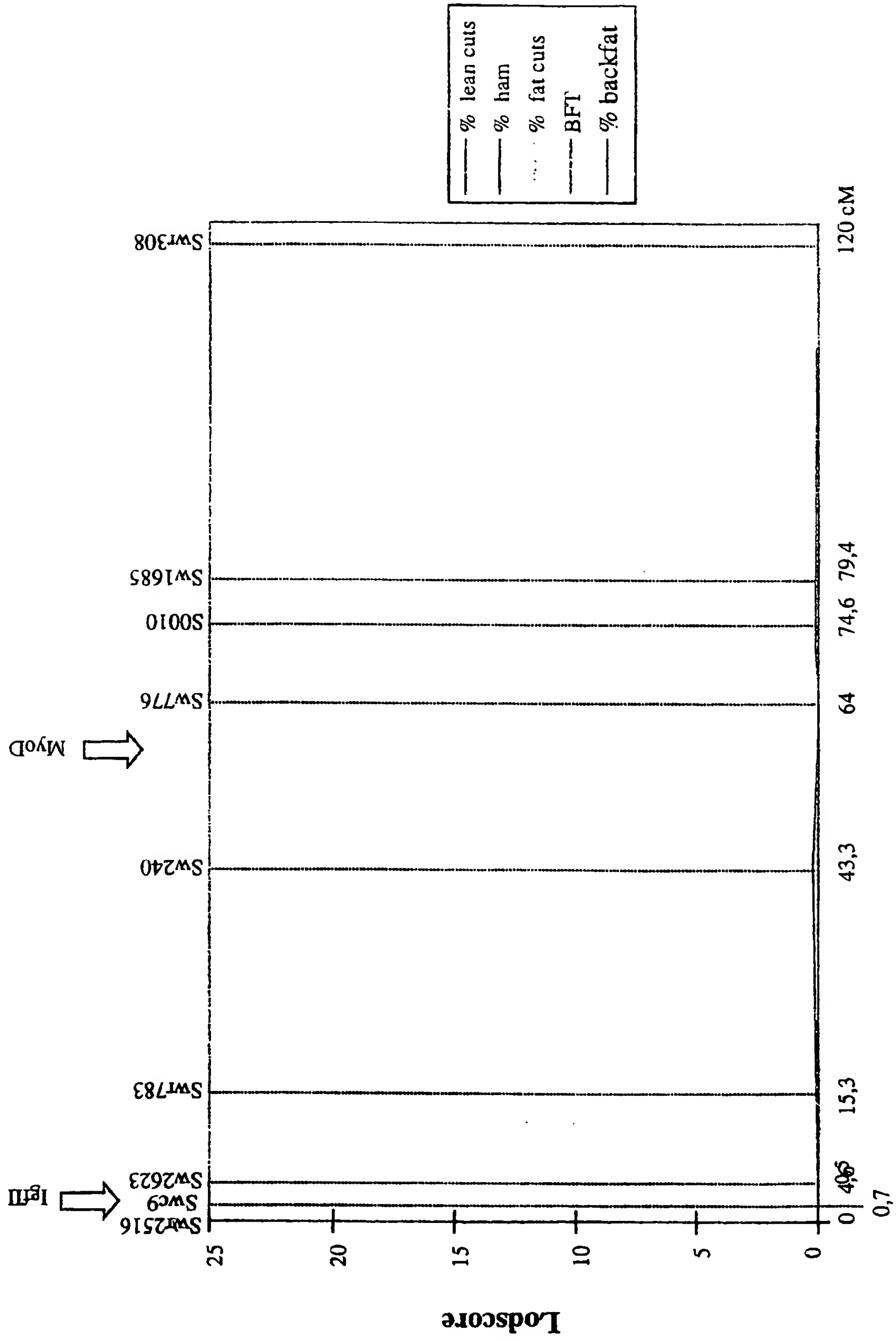
% lean cuts
% ham
% fat cuts
BFT
% backfat
Sw2516
Swc9
Sw2623
Sw783
Sw240
Sw776
S0010
Sw1685
Sw308
IgfII
MyoD
0
0,7
4,6
15,3
43,3
64
74,6
79,4
120 cM
0
5
10
15
20
25
Lodscore

FIGURE 4

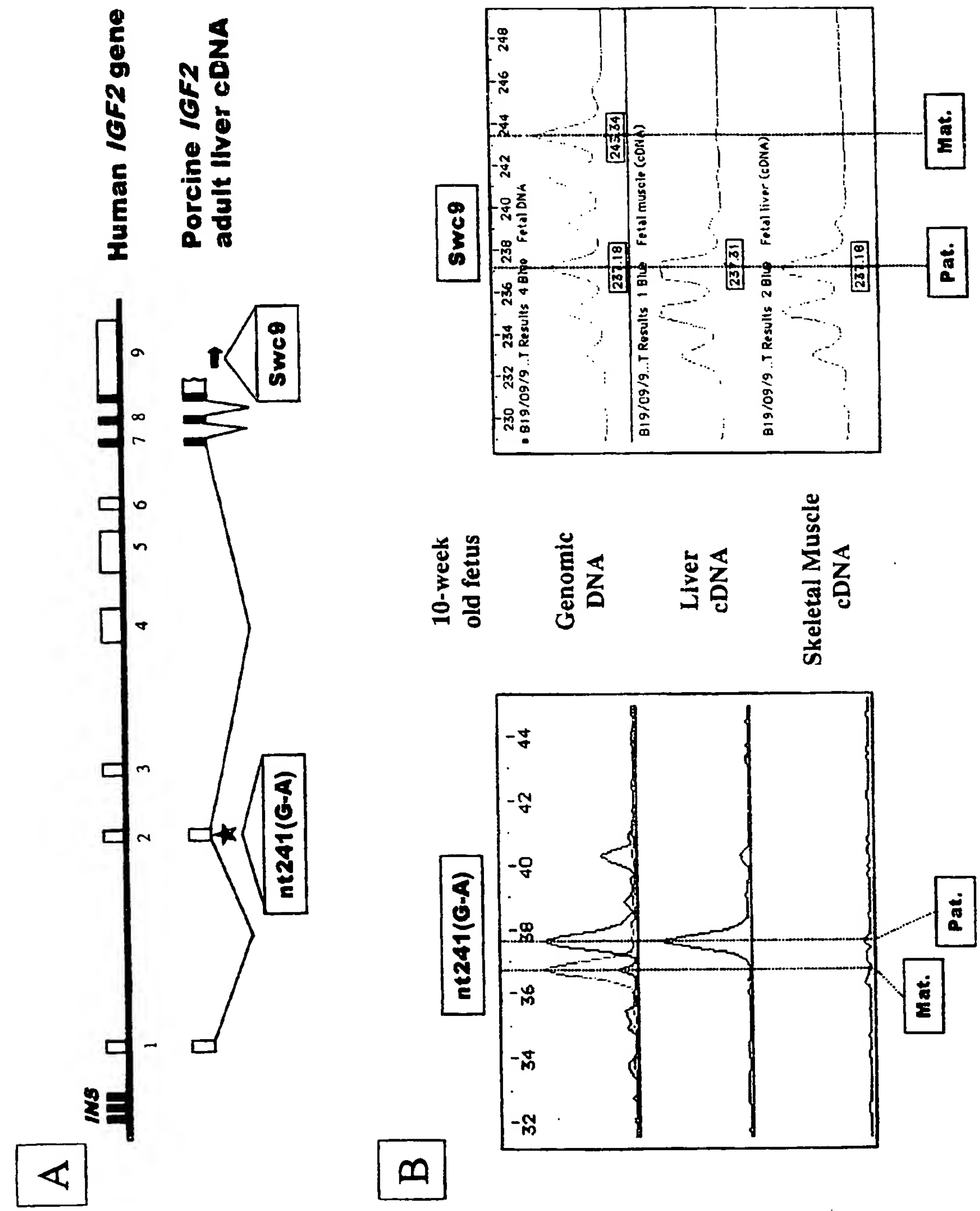
A
Human IGF2 gene
Porcine IGF2 adult liver cDNA
INS
Swc9
nt241(G-A)
B
10-week old fetus
Genomic DNA
Liver cDNA
Skeletal Muscle cDNA
Pat.
Mat.

FIGURE 5

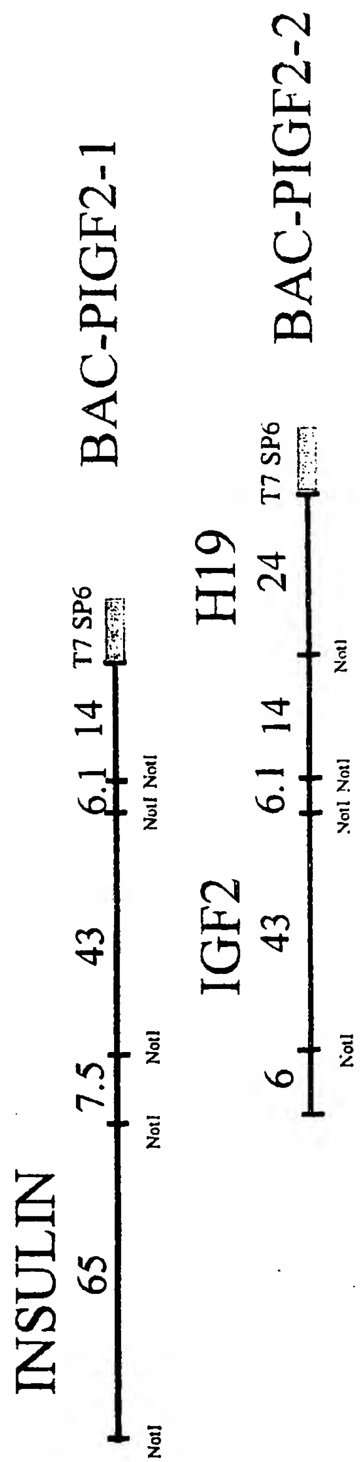
BAC-PIGF2-1
BAC-PIGF2-2
INSULIN
IGF2
H19
65
7.5
43
6.1
14
6
24
T7 SP6
NotI

FIGURE 6

Contig 1 (500 bp)
GGGTGGGCAGCTTCCTCCCAGACCGCAGGAGGCCCAAGTTCCCTGGCCCTGCCCACCCAGGGCCAGCTGAAGC
AGGTCAGAGACACCCGCTCCTGTCCCTCCTGTCACCTAACCCAACAGGCCGGGGCCCAGGGACACAGGCCACA
TGGCATCTCCCCCCATGCCCCTGCCCCAAGGCGCCCAGCAGGTGAGGCTGGAGCAGAGTCTGGGTCCTGCGGG
CCAGACCGAGGGCAGGACAGCTGGGCATCTGTCCTCACAG**TCCCCGCGCTTTGTCGGAGGCGGCAGAGCCTC**
ATCCAAGACGCCCGCAAGGAACGGAGAAGGCGGAGGCCGCGGCTGCCGCGTCCGAGCCCGGGGAGGCCCTGG
AAGTGGGGGCCCTTGCCGAGCGGGACGGGAAGGCCCTGCTGAACCTGCTCTTCACCCTGAGGGCCACCAAGCC
CCCCTCGCTGTTCCGGTCCCTGAAAAAATTCTAGGTGAGGGGGCGGGCCAGGGCTCCCCGGG
Contig 2 (943 bp)
TGCTCCTCACACCCCGGGCGGGGCTGCTCTTGGGGCCATCCTCCCCATGGGCCCAGCACCCACTCTGGCCTTC
ACACCTGCCGTCTTCTGGGAAG**TCCTCTGGTTCCCAAGGAAAGTTTCTGAGCTGGACAAGTGCCACCACCTGG**
TCACCAAGTTCGATCCTGAGCTGGACCTGGACCACCCGGTGAGCCGGTGCCTCCCCTCCCCGGCCGCCATGTC
TCCCATCCCCAGGGGTGTCCCCACACTCAGGGCCGGGACTGGGCGTGAACCCCGGGTTGGGACGGATGTTGGC
CTGCTGTGTGGCTCCTGGCGGAACAGAGAGGCCTGGCTGGGTGCCACCCCCAGGGCCCCCGCCGATGACACGG
GCCGCGTCTGGGCTGGGCGGGCAGGGCGGCCAGGC
AGGGCAGCCTCCGATGGCGTCCCCGGCTGTCACCAG**GGCTTCTCGGACCAGTTGTACCGCCAGCGCAGGAAGC**
TGATTGCCCAGATCGCCTTCCAGTACAGGCAGTAAGTCCCTCCAGGGCCTCAGCCTGGGGGCCCAGACCTCAG
CCTGGGCCTCACGCCAGACCTGGGGGTGGAGGGAAGGGAGGTTGTCTTTGTCACCAACGCCACCACCTTCACT
GTCACCATGGTCACCGACTCTGGGTCCCCAAATCACAGCTGAGGAAACTGGGGCACAGAGTGGTTAAGCATCT
TGCTGAAGCCACACAGCTGGCGGAGCATTTGGCCCCGGCCCCTCCTGCGGCTCCCACACGTGCTCCCTGAGGG
GCCCGGGACTGACAGCTGTCCCCTCCTCAG**AGGTG**
ACCCTATTCCCCGCGTGGAGTACACAGCCGAGGAGATTGCCACCTGGTGAGGCCCTGTGACAGCCGCTGGGAG
GGGCGGGAGTGGGGGAAGGGACAGGAAGACCTCAGAATTCCCGCGTGGAACGTGGTGGCCTCTATCATGA
Contig 3 (1500 bp)
GGGGAGGGGATGCTCAGACCCCGCTCTGGGAAGAAGAGAGCCTCAGAAGAAATCCCTTCCCAAGGGTCACGCGG
TGGAGCCCAGGGGCCCGCTAGGGGCCGATTCCCACAGCTCGTGCTGCCACCTGCTGGCGCTCCCAGGAACTGC
GGAGGCGGTGGGGGCCCTGGATGGGTCCGGCAGTGGGCTCGCAGGAGACCCCTGGAGGGGCTGCGGACACCCC
AGCTGCCACTCACAAGGTGCCCAAGCGGCGGTGGCAATGGGCTGAGCCTCTCCCCCCCTCCTCCTCCGCAG**GA**
CATTGGCCTCGCATCCCTGGGGGTCTCGGACGAGGAAATTGAGAAGCTGTCCACGGTGGTTTCTCCCCCTGC
AGGGCCCTGGGTTCCAGCCAGGCCCTCCTGTCCAA
GGGGTGTCGTCCTCACGCTGTGACCGCCCGGGAGCCTGGATCGGTTCTGCCTGGGTGGGCGGTGCCCGGGCCA
CGGGCAGCAGGGGCAGCGGTGCGGGCCCCAGCCGTGTCTGAGCCCCCTTGCCGCCTGTCCCCACCAG**CTGTAC**
TGGTTCACGGTGGAGTTTGGGCTCTGCAAACAGAACGGCGAGGTGAAGGCCTACGGGGCTGGGCTGCTGTCCT
CCTACGGGGAGCTCCTGGTGAGGCCTCCCCCACGCGCTGGGGCCTGGGTCCCCGGGGGAGGTGACCCCTGCGG
TGCCTTGTGGATTCCAGCTCTCGGGAGGCTGGAGCGAGGGGCTGCCCTCCTGGGGGCACCAAGAAAGCTGGTC
TGCGCCCCTCTCCACACACCTGTGCCTGGGCCCTG
GGGGGACCCCTGCTGGGGGATGTGGGTGCACAGCCAGGGCCACCAGGGAGTCAGGACACGGGGCTCCCTTCCC
TCGGGTCCCTGAGACCCCTGGCCTCCCGCCAG**CACTCCCTGTCCGAGGAGCCCGAGATCCGGGCCTTCGACCC**
CGACGCGGCGGCCGTGCAGCCCTACCAGGACCAGACCTACCAGCCCGTCTACTTCGTGTCTGAGAGTTTCAGT
GACGCCAAGGACAAGCTCAGGTGGCCGGGGCCCGGGGCCCCAAACTGGAGGATCCAGCCTGCAGCCCCGCC
TATGAGCCCATTTCCCAGCAGAGGGAGCTGCTGCGGACCCCACCGTCACAACCCCCTCCCACAGCTGGAACC
CCAGAAAGCCTGCGGAGGGGGGACCTGCAGGGCTG
TGGCCAGGTCAAGGCCAGGTCGAGGCCAGGCTTTTAGGGGTGAAGTCTGACTTTGTAAGAGGGGGTGCAGGGT
CCTTCCCAGCCTCCTCCCCTCCGAGCAGCTGGGGGCGGGGCGGGGGTGCGATGAAGGCAGAGATGACGCAGCC
ACCCGTTCACCCTCAGGAGGCGCCTCCTGTCCAGCCAGGCTCCTGTTGTCACAGGGGAAACTGAGGCCCCAGG
TGTGTGTGTGGGGGGGTGATTCTCACACACAAGCTTAGGGACAGGGACATAACGGCCTCTCCAGGGCACACAG
TCTGGAGG
Contig 4 (3024 bp)
TTAANTCCANGTTGGCCCGACAAGTTTTCCCCATTTGAAAAGGGGCCAGTTAAGCCCCAACNCAATTAATTGG
AAGTTAGCTCCCCTCATTAGGCTCCCCAGNCTTTACNCTTTATGTTCCGGTTCGTATTTTTGTGGAATTGTA
GCGGATACAATTTCTCTCAAGNAACCAGCTATGCCCATGATTACGCGGTACAGTAGTTCATCAGTCCCCCCG
CCCATGGGACAGCGAAGGGAACCAGTATGTCGTGGGGCCGGGTCTAAAGGGGTCACCACCAGGGAGGGGCAGG
GGCTCCAGGAGGCAGGGCCACTGAGCGGTACCTGGTGGGGGGAGGTGGTGGGGCCACACCCAGGAGTCCTGTG
CCCCCCCCACTCCCGCCGTTGGACATGAGAAGCAGGGGCCAGCCTGCGGGTCCCTGAGTTCAGCGCCCCCCCC
CCCCACCGCCGCAGCAGCCCGGGGTCTCAGCAGGCTGCTGTGCTGGGGGCGGGGGCGCTTATGGRGCCGGGAG
CAGCCCCCCCCCACGGCTTCAGAGCATCTCTGGGGCCTCAGGGATGGACCGGGGTCTGCRGGCAGGTGTCCTC
TCGCGCCCCCACTCCCTGGGCTATAACGTGGAAGATGCGGCCCAAGCCCGGKCGGTTTGGCCTTTGTCCCCAG
CCAGTGGGGACAGCCTGGCCCTCAGGCCGCTCGTTAAGACTCTAATGACCTCAAGGCCCCCAGAGGCGCTGAT
GACCCACGGAGATGATCCCGCAGGCCTGGCAGCAGGGAAATGATCCAGAAAGTGCCACCTCAGCCCCCAGCCA

FIGURE 6, CONTD.

TCTGCCACCCACCTGGAGCCCCTCAGGGGCCGGGCGCCGGGGGGCAGGCGCTATAAAGCCGGCCGGGCCCAGC
CGCCCCCAG**CCCTCTGGACCAGCTGTGTTCCCAGGCCACCGGCAAGCAG**GTCTGTCCCCCTGGGCTCCCGTC
AGCTGGGTCTGGGCTGTCCTGCTGGGGCCAGGGCATCTCGGCAGGAGGACGTGGGCTCCTCTCTCGGAGCCCT
TGGGGGGTGAGGCTGGTGGGGGCTGCAGGTGCCCCTGGCTGGCCTCAACGCCGCCCGTCCCCCAG**GTCCTCAC**
CCCCCGCCATGGCCCTGTGGACGCGCCTCCTGCCCCTGCTGGCCCTGCTGGCSCTCTGGGCGCCCGCCCCGGC
CCAGGCCTTCGTGAACCAGCACCTGTGCGGCTCCCACCTGGTGGAGGCGCTGTACCTGGTGTGCGGGGAGCGC
GGCTTCTTCTACACGCCCAAGGCCCGTCGGGAGGCGGAGAACCCTCAGGGTGAGCCGAGGGGGYGTCCCGGGA
GCGGTYGGGGGAGTTTTTAAGGAGGAAATTGGTAAAAGTGACCAACTCCCTGGGAGCTGAGCCCAGAGACACC
CCTCCCACGCCCYGGTCCCGCTCGAGAAGCCCCCCTTCCCTCCCCTCCTCCCG
AGGCGGCTCCAGGGAGGAATCTTACGGAGTCAAGGCCCGGGTGCCGCTGGTCTCCCAGTGACATGGCCGTGGT
GTCCCRTCTGCCGGCCCACATGCCCGTGAGAGAWGCCCCATCCCCCTGGCAGGGGGCCCCGTGCCGGGCAGGC
GGCGGGAGGCCCAGGACCGGTGGCTGCTGCGGCTTCCACTCCAGGGTGGGCGGGGTGGGGGGTGGCTGTCTCT
GTGTGACCGGCTCTCCCCGCAG**CAGGTGCCGTGGAGCTGGGCGGAGGCCTGGGCGGCCTGCAGGCCCTGGCGC**
TGGAGGGGCCCCCGCAGAAGCGTGGCATCGTGGAGCAGTGCTGCACCAGCATCTGTTCCCTCTACCAGCTGGA
GAACTACTGCAACTAGGCCGCCCCTGAGGGCGCCTGCTGCTCCCCGCACCCCAAAACCCAATAAAGTCCTGAA
TGAGCCCGGCCGAGTCCTGTGGTCTGTGTGGCCTGGGCGGGGGCCCTGGTGGGGGAGGGGCCAGAAGGCTGT
GGGGGGCCTGCCTGCGACCCCTCTCTGCTCTCGCCACATCGGCTGCTCTAAGCTTCCTCCACATGCATCGGGT
GCCCACAGGCACATGGGCACCGGGGGACCAGGGCCCAGGGCAGGGCCCTTCAATGTGGCGAGCTCTGGTTTTC
AGGGCTCCAGACACCCCCTCCTGGGTGCCCACTGCTGCACAGGGTCACTCTGAGGGTCACAGGGCACCCACCC
AGACTGCTCTTGGGCACACAAAATAGCCCAGGGGCTTCTTGGGCTGGCTGCRGTCTGGGAGGTCAGAGAGTGA
CCCCGCGGGACCAAGACCTGGCCAGCCTGCCAGTCGCCCAGGCCAAACCAATCTGCACCTTTGCTGAAGGTTC
CACCCGGGCCAGCACTGGGGGCGGCCGGGCCTAGAGCTGGGCGCCCGGGCCCCAGGGACTGCACACCCGCCAG
AGGTGGGCCTGAGGGGTGGCAGCAGGCTCTCCGCCTGGGACCCAGCCAGCTGGGCAGCTCACCTCTCAACACG
AGGCTCTCACCTGTGTCGTCCCCTCCCCACGGCCACACAGACACCCCTGGGGAGAAGTCACAGGCCCCCAGCA
GGCCCCGCCCCTGGAGAGGAGGCCAGGGCTGGGCAG**GCGGGTGGCCGGCCGGACACTGGACCCGGAAGGGGGG**
TAGGCGGCTGGGATGAGTGGCGAGCTGTCCATGGGAGCACCCAGCGGCCCCATTGGCACCAGTACAGGCAGGG
GCACCTGCAGCAGCTGAGGTACGTGGGGTCCCCGGACTGGTTGGTGTCCGGCTGCCCTCTGGGAGGCAGCGGG
CTGAGCTTGTGGTCCTGCCAACCAGGGAGACCCGTGACCACCCTCCTGCTTCCCCTCCCCCCCAGGGCCAGCA
GACTCCTTTGGGACTCGGGGCCCCTGAGCCGCCCCCACTCGCAGGACTCACGGGGTGTGCGGTCCTGGGTGAG
TGGGGGCTTGGGAGAGGGTCACTCTTGTCCGTCGGGTGGGGAAGGCTGAGAGTCATGGTGTGACAGCGCCCTC
GGCCTGCCGGGTGGGGGGTCTCCCTTCTCCCGAGCCCAGATCCCCGGGTAC
Contig 5 (1730 bp)
CGTCACCCGCAGAAGCCAGGCCCACAGGCCTTGGCTCAGCCCCTCCACCCAGGCCCACGTTCCGCCCCTTCTG
GGAACTGGAGGACAGCCCGCCCTCGCCCTCGGACCTGGCTTCGTTTGCCCTGGCATCTGGCAGTGGCCGGCAG
CTGCGTTCAGCCCTGGATGACACCCTGGCGTGAGCGGTGGGTCCCCGTGCTGAGGGCAGCCCCCACACACGTC
CTGCTCACTTGCCTTGTGTCTGCTCCGCATCCCGTCATCACACATGCCATGCTGGGGCACCGTAGCGCCTTGC
CCTGTGTGGCACTGTGGCACTGTGTTCCTGATGGGAAGACTGAGGCTGGGGTCAGGCCCGCTGCTGCCCACCC
TCTAAGGACATTCTGCCGGTGCAGCTGCCTCCAGG
CTGGCCCCCCGGATTGCATCTGCTTCTGGCACGGATGAACTGGCACCTCTGCCTGACCATTAGGGCTGTATTT
GCCTTCTCCTGTTGGCAGTAAATATTTACTGTCCCTCCCTGTTCCTCCAGGCCCGANCCAGTTCCTGAGGGGC
ATGGGAGGTGGACACAAAGGTGCCCAAGCAGCCCCCTGCTCTTGAGGGCCCAGTGTCTGGTGGGGGCCAGCCT
GGGAAGGAGGAGCCAGACTAGGAACCAGAGGCCTGTGTTCCTGGAAAAGGCCCCCTGGCAGAGTTCCGGCTGG
TGTGTGTCCAGCTAGGCTGTGAGTCTTCAAACTGGGGAGCCCGGCCCCTGGACCCAGGCAGGGCTGCACCCCT
GGTGCCAGTGCTTCACTGGGTGGGCACCTGTCCCC
ACCAGGCAAGGTGGTCCGAGCGGTCATTCACAGACAGAACCAGCAGAGGGCGCCAAAGCCCCACTTTTGACAA
ACTCCCCTTCGCCCTGAGCCGAAAGTCCAGGCGGCAGGTGGACCTCTCTGCAGGGCTCTGCCACCCCTGCTGC
CGCTTGCCAGCACTCACAGGGGCTGCGGGGGGTGCCCAACAGGCCGGCTACCCTGAGCTCTGGAGGCGATGGA
GTTTAGGAGGGAACGAGGGGACTCCTGGGGGTGACTTTCTTCAGCGCCCACATTGCGGCCCAGCAAACCGAGG
CTGGAGGAGGCCGGGCACCTGTGCCCAGCTGGAGCCTTTGCTGAGGGTCTCCAAGGCCTGGGGAAATTGAGGC
TGGGGGCTGGGGGGTGTCACTGTCGGGCCAGGAGG
CCCCTCGCTCTGATTGGAGCCGCCTCGGCCACTTGAGCCAGGAGGCTCACATGAGGCGGGGGCTGCAGGGACA
GGACCCTCGGGGCCCGGGAGGCCTTGGAGGGGGTCCAGCTGGGCCAGGGTTCGTTCTTTCCCGGGTCCATGTC
CACCGCCCTCCCGCTGCTGGGAGGAGAGGAGGTCCAGGGCAGAAAGAATGCGTGGGGATGGGGGGGTGGTCAG
GGGTCTGGGAGCTGTGGAAACAACAAACAGACAGCGAGGTCCTGGGGCGCCCGGCCCCCCGCCCCCTCCGGCA
CTGTTGTTTCTGGCCGGGCTGCAGGGACAGCGAGGCAGATTCCTTCGAAAGTGGAGACTGGCGGGGGGCCCCT
CGGGTCCTCAGCTCACCCCCTGAGCTAGCCCGCCC
ACTCGGCTCCAACCTCCCGCAGGCCCCTGGCACGGTCTCCAGGAGTCCACTGAGGGGTCCCCAAAGCTGCCAC
CAGGAGCTGGGCCTGGGTCTGTCACCACCCCACCCCACCCTCCAAGTCTGAGATATG
Contig 6 (4833 bp)
ATGTGAGCTGCACAGCATGAGCCCTCGGCCCCACTGCTGTGGCCTTGCGGACATTGAGGTGTGTGCCGCCCAG
GGCGACCACACCCTGGCCTCTCAGGGTGCCCGTACAGAGGCGGCTGGGTCGTANGAGGTGCGGGGCTCTGGGG
ACCGCTGGTGAGTTCAGGACGGGGGTCATGCCACCTCCTCTCTGAAGGTTTGGTGAGGTGGCCCTTCTCTTAT
CGTGATGACAATACTGATTTCTGGAAGAGCCAGGTGTTTTCTGAGGCTGTGGTTGCACTTCTCCACGTGGCCA
CAAGGTGCCGGGCTCGGGTCAGATTTGAGAAGCCCTGCGGGAGCGGGTGTCATGCGCCAGATTCAGCTTGCCT

FIGURE 6, CONTD.

CCTGCGGGTCTGGGGTCAGGACCTGGTCCCCAGCAGTCTGCTCCAGAGCCTGTCAGTGATGTGTGGGATTTTA
CCGCTAGAACACAGTTTCCTCTGATTCTCAGAAACCAGCAGATCCTTTAGGAGGGGCGTGCAGGTTTCACCTG
TGCTGCANNGCCCCCTGCCACCTGGTCGGAGCCNCAAGACGGCATCTAAAGATCAGTTCCTCATCATCAGTTC
CGCAGTGCTGGGGTGGGGCAGATGAGAACCTCAGGGCTGGGCGCAGAGGTGGGGAGCCCGCCTGGACCCCGA
CACTGCAGGGGGGCCTCCCCCTTGTAGGAAGAACAATGTCGCTTTGCCACCCAGCCCTCTCCCCAGGGTGCCC
CGAACTGTTGCTCCTAAGACCTCTGGGCTGTGTGCTGTAATTCTATAAGTGGCCACCAGGTGTCAGCAGGAGG
CCACTTAAGCATCCATGTGGCGGAAACCTGGAGCTGGGGGTTCCTAAGGGTCCCTCGAGTGTCTCCTGAATAA
ATAGGCGCTGACCTGATCCCCAGGAAGGGATAACCCTCTCCCAGGCCTAAGAGGCAGTGGGGCAATGAGGTTT
ATGTGTCCACTGTACCCCCAAATTGTCTCTTCCTTCCCTCTACCCTGTGTCCCCACCGTGGACGATACACGGA
GTGCGAGGCTGCGGGTCACAGCCCTCACAGCCCCAAAGCTGCAGGTCCTGCCTCAGGGGCACCGCAGCTTGGC
TGGTCCCCCTTGGGTCCTCCCCACCCTGACCCGTCCTCTGCTCCCCTCCCTTTGCTTAAATGCTCTGCGTTTC
AAGGTTCTGATGGAATAAAATAGCCCTGCACTGGTGTGTTCCTCTTTGGGGCTGTGCCAGAAGTGGGAATTCA
GACCAGGGCAGAGCTCAGATTCCACATACTGTGTTAGGGATGGCAGGTGCCACATTTCCAGGAGTTTCATTGG
TGGTTTGTAAATGCTACTTCCGTTTCAGCCCCTCAGCTGCCCACCTCCTCAATTTAGGGACCCCCCCCTTTGG
CGGGTTGCCCATGGAACCACATCATCTGGCGTGGGGTGAGCCCTTTATCCTCCCTGGCCCCACTGGGAGGGTT
TGGGGAAGTCCCAGCTAAATTTCTCCGTAGGGACCTGGAAGGAGCCCTTGTGACATCTGGGCACAGATAAGAG
GTAGGGGGCACAGGCCGTGAACACTTGAAGCTGCAGAGCCCAGAGCAGAGCCAGCAGGAGCAAGTGACTGCTC
CCCACCCCAAGAACTGTGGGCTGCGTCACACACTCCCCACTGTGTGCCCTGGACCTGACAGGGCCTTTAGCCT
CCCTGCATCCCTCCCCACCCAAGAACCCAGTGAGGCACCCCACTTGCCCCTCCTTAGTGTTGTTATGGCTCTG
GGGCATCTGCATTTTGTTTAGGACACCCCCAGCTAGATTTAAGTCCCCCCAAGTGTGACTCTTCCTCCACTG
AAAACCCTGTCCTCCCACCAAAGGGCCCTATCCCTTTAGCTGAGCCAAGGAAATTCAGGAGGGGCCTTGAATG
ACAAAGGAAGAGGGGGAGAGTTAAACCCCAACACTGGCTGGCAAGCTGGGTGGGGTGGACACCCCAGGGTGCA
GGGGTGCAGTGAAGGTAGCGGCTGGTGGCCTTCTGGAAACTACATGTGACTTTGCCATTAGGTGAGTCTTTGC
TTTGCCCCTGCTCTATCTGCAGGCTTATGGAAGAAGTTTAAATTCCCAGGGACACTTGGTCTAACCAGGCAGC
GCTTGTATCTGGGCCCTTCCCCAGCTGCTGACCACTCTGAGTCTGCGCCTTAGTTGGAGTTTTGGCCAAGCTC
AAGAGGCTGTGGACCCCAGTCATCCCACCCAGGGGTGCCTGTGGGCAGGACGCTGCTGCCTGCCATTTGCTGC
AGTATTGTCACTGTCCGGCACCACACACATGGTGCAGGGGGTGGTATCAGGTGCCACTGGGGAAGGGAGAAAA
CTCCCAGGTGAGTCCCCTGCCTCTGGAAGCAAGATGGACATGACCGCACTGTGTTGCAGCTGCATTGGGAGGC
CCCGAAGAAAGATTTTTCTGATCTTTCTCGAACCCTGCTTTTCCCCATCATGCCCCGCCCCCATTTTACCCGT
GCCACGCCCACTGGTGTGCCGGGGTGTCAAGTGACTGACAAGTGTCAATCTACTGAGGCCCTGCCCACTCTCC
ACCCCCCCACATAGTCCCACCTCCCAGCTGGCAGGGAGAACTTCCAGCTAATGCCCATGCCCACAAATGTCTT
TCTGTCAGCCTAGAGCTGGACCAAATCTCCACCCTGTAACATGCTGTGCCCTGGCGTGGGAAGGTGCCAGAGC
CAGTTGCCCCAGCAGCCCCAGAACCACTAAGTTGGCACAAAGCTACCCAAATTTGGAGGGGCTTGGGGAAGGG
CATGGAGGGGATGAGGAGGTGAGGGGCAAAACTAATTTCAGTTAGCATTTGAGCAGGTGCCACGCTCAGCGTG
GAGAGGCTCTCTTGCTTCTAGGGACCCATTATGATGCACACGCTAAAAGCGCCCTTCACCATCTCTCCAGCCT
CAGCTTTGTCCCCCTCCTCCTCCTCAGCGGCAACCCGGCTGGAGGGTCTGGCCACTACAGCCAGAGCGCCCCC
TACTTTGGTGGCGACTGCTACTATTGGCCCAACCAGCGGATCACCGGCCAGGCAGTTTCGGCAGAGAGTCTGG
GGCACCAGTGACTCCCCCGTCCTCTTTATCCACCACCCAGGAGCTTCAGGGACTACACAGCGACTAGAGGGCA
GGTAACTGGTCTGCCCTCCCTAGGGCTGCCCCCTCAGAGTGTGTGAGAAAAGCTGCATTGAGTGTTTGGGTGC
AGGTGGGCTGGGGGCTTGGGGCAGCCAACAGGAACGGCGGGACCTCTGCTTCCAGAGGACCCCAGATCCTGGC
AAGCTTCGACTTTGGAGGGGACAGGAAAGACAGGTGGAGAGGGGACACTTCCCTCTTCTGTACAGACGCCCAC
CCGGAGCCACAGAGGCTTTTGCAAGGAAAATAGGTTTTCCCTCACTAATGCAGCAGGCAAAATGGGAGGGGCA
GGGGTGGAGGGTAGTGCCCCCGCCCCCAGCAGGAGGGCACAGCTGTTTCTGCAAATGTAAAAAAGCAGGGTTT
TTCTGTGTGAGAAGTTCCCTCTTGCTGCATGTCCCCACCCCCGCCACCAAAGACAAACAGGACACTGTGCAGA
GGGGCCAGAGCCCCGAGATTTTGGACTTGTTTTTATATGCATATATACCATTTTGAAAGCAAAGCTTCCCTCT
CCCCTACTCCCTACATGTCCCCCTTCACCAAAAAATCCCACCACGTAACTGGAAAGGGGAGTGAGAAGGACGA
CGAAGGGGCACTGTCCCCTCCCGTCCCACAGCGGGACTTAAAACGTACAGCTTTTCGCCTCCGGACAGTGTGC
CGCCCCCTGGCCCCCGTCACGCTCCCCTGCCCGGGGGGCTGAGTGTGGGGCCAGGGCCTGTCTCCAGGCATGC
ATTATTTTGTGCATGAAGGTTTTGTCCCGCCCACCCAGGCTGGTGTTGGGGGGAAGGGGTTCATTGCTCCAAA
GAAGCCCATCTCCCCCCTCAGCCACCTTCAGCCGCCTTCGCAAGGCAGAGCTGTGTCCTCTGCTGTGTGCCTG
GCCCCCTCCTTGCTTCTATTCAAGGTGGAAGTGTTGGGGGGAGGAGAAGAGTTTTTATATTGTGTCTGTGATC
CCCCGAGGCAGGGCATTTGTGTGCGGCCCCCCAGCCCCCAGGCCCAGGCAGATGGGCCAGCCTGCCCGACAGA
AGGGTCTCCTGCTGCTTGGCTGCAGGGAAACCCAGCTCTGGGTGAACCGTGGGCACCTTCCTTCCTCCATGCC
CTGTATTTAAAGAAGGAGAGCTGGGGGGCCAGAGGCACAGGGAGGGGAGCCACGGCCCCAGGTCTGACAAGAT
GACCTGCGGGCCTCTCCACCCAAGAGTCGGGGTGGGGGGGCGGATTTGGTTTGAAAAGAGAACAAATAGGAAC
ACACTCTTTATTTTCCCCAG**GGGCCGAAGAGTCACCCCTGAACTTGAGGACGAGCAGCCGGATTCCAGCCCCC**
AGCCCCAGGGCCCCACATCTCCTCGGGCTCAGCCGCGCGCCCCAGCTGCCCCCCAGCCTGAGCTGCAGCAGGC
CAGGGCTGCCCGAGACCCCAGCCCCCAGGTGAGCTGCTGCAGCCTGTGGCCCAGGAGATCTCCGCCGGCTCAG
AACTGAGGCGGGCAGCCCACCCAGCCCACAGCGGTGAGTGTCTCCAGACCCCAGGGCAGGGCCCGGTGTCCCC
CGGCACAGAGAGCTGTGCTGCAGCCCCAGACCTCCCAGGCCGTTTTAGTTCCCATCTCCCCTTGGGGGAGGGG
TGGGGCTCAGAGGGGCTGGGGTGCATCCGCAGAGCTGGGGTGCAGGGCTCCAGGTGCCTCTCTCCCAGGCGGC
TGGCCCGGAGGGGGG
Contig 7 (2014 bp)

FIGURE 6, CONTD.

CTGGTTTCGCACTCCTCCGGGGACTGTTGAAGTACCCGAGAGCGCNCGCGGAGCGCCGGGGCGAGCGGGGGTG
GCCGCCGGGGGTGCTCCCGGGCCCCCGGACCGAGCCAGGGACGAGCCTGCCCGCGGCGGCAGCCGGGCCGCGG
CTTCGCCTAGGCTCACAGCGCGGGAGCGCGTGGGGCGCGGCCGCTGCCGGGAGTCCGCCTGCCTCCTCGGAGG
CGGCCGACCGGGGAGCCTGGGGGACCCCGAGCGCCCGGGGAGCAGCGCCCCGACACGCCCCGGGCCGCTCTCG
GCTTCCTCCCTTCCAGCCGGCGCCCGCGCGGCCGGGCTTCGGCACCGGGGCGCTCTCAGTGGCAGGAGAAGCG
TGCGCTCCCGCGGGGTGGGGGACCCGCAGGAAACC
CGCACCGCCTGGAGCCGCCGCCGCGCGGCCAGCGCTCGCGTCCCCCGGGGAGGGCGCCACTGCTCCGCGCGCG
CGTCCCCCGACGCCCCGCGCGCTTCCCCGGCCGGCCCGGGATCCTAACCTCTCTCTCGGTCGCAGCCCCGCAT
CCCCAGGGCTCCAGGCCCCCGGCGACTTGCCCGCTCCTCCCAATTGCAGACACGACTTTTTCTGGACCTCCC
AAAGGACAGCCTGGCTCCAGGGTCCCCCAGATACATTCACCATTTCTCCAGATCACAAGTGGGTTTTTCGGGC
ACTAACTTCCAGAGACCTCAAAGCACATGAGCCCCTACTGGCTTTCCCAGGTTTCCACTAGTGGCCTCGGTCC
CCACCTCACTGGGGATTGTCTCCCAGGCTCTTCCC
GGTGTGATCCCACCCATTCGCGCCCAGGTCCCGCAGTGCCAATCCCTCCTCTAGAAAACTTAAACACTGACTC
CTGGTCTCGGGGTGAGGCTGCCCAATGTGCCTGACTCCCCAGAAGGTATACCAGTGTTTTTCTGGCATTTGGG
CACCGTTCCCCCAAAACACGTGAAGCTCTTTTCCCGCGTCCCCATAATTTTGGACGCCAGGGGCACCCAAGCT
TAGCGCCCCTGTTTGGCTCCCCCACACCGCGAAGCCCTGCTCCCTGGGGTTCACGACAGTTTGGGACTTTATC
TGCCAAGTTCCACAAACTGATTGGCCCCAAGCTGGGGTCCCTAAATTGTACACAAAGAACCCCAGCCCCCCCC
CCCAACTCCAGTACAGGAAGCGATGGCCCCAGGGA
CCCTCGGAGTTGGAACGTGGCTTCCTAAGCCTTCACCAAAATTGAGGCTTTCCGCGCATGGCGCGCTGATGCC
CTTGCTGAATCAGAAGCACTCTGCCCTCTGATTCCTGCTTTCCACAACCCTGAGAGCATGATTTCTGGTCCCC
CAAACTCACTGAGCAAAAAATCTTTTTGTGGGGGCTGCAAAGATAGGAGGCATTTCTCTCCGGAGCTCTCCAAA
CTCCCTTGCCTATAATCAAGTTCCCTAAAACTTAGACAGAGCTTCCCAGGCCCCAGAGGCACACAGAGCCATT
ATTGGAGCTGCGTTTAATGATGACAGGGACCATGGGTCATGCAGCTCCCCCAAGTCACAAATGCCCCAGGTAT
CCTTGGCTCCAGCCAAGCCCAAAGCAAACTCTTGC
ACAGATCCCATATCTTGTTATGTCAAGCGCTTTGCGTGTCCCAGTAAACAAATAGTCTGAGTGTTTTCTCCAC
CTCATAACATTCGGAATATTAAAAAATTCCCTGGGCCCCCGGAGCTGACAGACAAGAATCCGGGCTTCCTAAA
ATTCAGAACTGATTCCCAAATCCCAGGCCAACGCCAGACCCTCTCCCAATCTGGAGCCCCTCCGACTGGACAC
ACTGGACTCCTAAGTATTACGCGCTGTCCTCCAGGCACCCCAAATGCATTCAAAGTGACGCTTTGGTCACAGA
AAGGCACTGATTTCTTGGGCTCCAAAGCAGCCCATGCACCCCCGAGTCACCCCAAACTTAGTCAGCATTTCCC
GGCTCTCCCTCCGCACTGCAAACTCCCAACTGCGG
ACACCGGTTCTTCAGGACCCACCGCCTAGACGGTCTTAATCCCTTTTCCCCCAGACCTAGATTC
Contig 8 (371 bp)
AGATTCAAAAACTATTTTTCTCGGGCCTCCAAATTGAGGTGCTGCCTGCCAGTCCTCCAAAATAAACTGAGGG
GTTTTTTGTTTGTTTGTTTTTTTGTTTGTTTGTTTTTTTTTTACCTTCCACGAAACAATCCAACTTTTTGGA
CCATTGATTTATGGGTCCCCTGACTTTATGACCCTTGCCCCAAGTCCCCCTAAATGTAGGCCATTTTCCACGG
GCCTCCCAAAATGAAATTGCCCAGATCCCGCCGAAAAAAATATCCCCGGGTCCTGGAAATCCCAGGTATTACA
GGCCTGCGGCTGACACCCCTCCTTGCTACTAACCAGGTTCCCTGAAGTTTAGAGATCACTACCTAATGAACAA
ATCCAC
Contig 9 (2415 bp)
CCAAAACTGGGGCCCTATCTTACTAGGGTTCCCTAAATGCAGACAGCGCCCGGGAAAATAGGGGCGTTTTTT
TCCTGTTTGCCAAAAATAAACTAATTGAAACCAATTTTTAGAATTAAAATCTAAAATGACCTTGATTTTCTGC
GTTCTCCAAATGTACTTTTCACAGCCCAGGTTGCCCCCAGTTTAGACGGTGTTGCTTGAATCTCTAAAGCACC
CTGAGGATTTTTCCCGAGGAAGCCACCACAACTACGGAATTTACTGTCCTTCGGGGCCACAAGCCTCCAGGCC
ACCAACTTGGATTTCTAAACCGTGGAAATCAGCCTCCACTTCCCTCCGCCACCCCGAGGGTCTGCTCAGACCC
CCCAAACGTGCCCGCTGTTCTTCTCCCCCCAAATT
TTATTTAGAGAATATGCCTCTCTCGGGTTCTGCCAAGTTTCCCGCTGAGACTTCCTCGGTCATCCCCAAATCC
TCTTCCCCACAGTCCGGGAGCCCCCACAAGCTTACCGACCCACATGCTGGGGTCCCCCAACTTAAACGCGATC
CCCTGTCCCCCAGATTCACCGAGTGATTTCCCTGGTCCTCAGACTGGGACTCTTTTACTGGAGTCTCGAATTT
AGCCATTAATCACAGTTCTCCACTCCGACGCAGGCTCCCTTGGGTCCCCACGTCGGGGACATGGGTTCTCTTG
CCTGCAAATCAGGCTGCTCTGACTTGCATTCAGGCCTTTGGGCATTGTTCCCCGCCCGCCGCGGTCTCGGTTC
TCCCCCCATCCCGCGCACGACGGGCACTGGGTCTG
GGCCTCTTGGTGTCTCCTACAAGTCCCCGGAGCTCCTCGGACTTGGGAACTGTCTCTTGCGTTCCCCAAATAC
ACTCGGCCCGGCAGTGTGTCCGCCAGGACGTAGGCAGAGCTTCTCCCGCGTCCAGGAAAACGACTGGGCATTC
CCCCCAGTTTCCCCCAAATTTGGGCATTGTCCCTGGGTCTTCCAACGGACTGGGCGTTGCCCCCGGACACTGC
GGACTGCCCCCGGGGTCTCGCTCACCTTCAGCGCGTCCACCGCCCGCTGCAGAGCGCTCGCTCTCCGTCTCTC
GGCTCCCAGCGCGCTTGGGGACGCAGCCTCCGGGCCTCCAGCCTTGCGGTGAGCTCCCCGTCGCCTCGCGTGT
CCCGCCCCGGCTCCCAAACCCACTCGCCGCCGTCC
CGCTGGGGCTGGCACTGGCCTCCGGCGACTGCCGGGGACACGGGAGCGGAGCGCGGGAGCCTGCTGCAGGCCA
GCCCGTCGGCCGGGCCGCGCGCCCTGAAACGCGCGCGGCTTTCGTTTGCTCTTTGCAAAGGTCACAACCGTGG
GGAAAACGCCTCGGCGGCCCCCAAGCGGGGCAGGCAGGGCGTTGGGAAGGAGGGACACGCGGGAGAGGAGCAC
CCCGCTGGGGCGGCGCAGCGCGGCGCCTCCAGCCGCCGGGCGGAGGATCCCGGGAGGCGCGCGCGGAGCGCGG
GCGAAGTGATTGATGGCGGAGCGAGGGGCCAGCGGATCGCGGGCTTCCGCCGGCGGCGGCCCCTTCCCCTCG
GAGGGACTCGGGCGGCCCGGGTTTCTGGGGGCGGG

FIGURE 6, CONTD.

CGGGGCGCGGGGGCTTGTGCGTGGTCTCCACTTGGTAAAATCACAACGACTTTTACGTCGCCCCGACTCTC
CAGGAGATGGTTTCCCCAGACCCCCAAATTATCGTGGTGGCCCCCGGGGCTGAACCCGCGTCTACGCAAGGCC
AACGCGCTGAGGACGGGGAACCATTATCCGGATATTTGGGTGGGCCCCCAAAGCGAGCTGCTTAGACGCGC
CCCGGTGAGCTCGGTCCTGCAGGTAGGCTTCGAGCGAGGTTCCCCGCCCTGCTCCTCTCTCTTCGGGCAGGCG
CGGCCAGGCCGGCCGGCCCTCCCCACGTACGGCACCTGGCGGCCGCCGAGACGACTCCCCGGTTCCCGCGCGG
CACCGGGGGGCGCTCGGGCTCTGGCTGCGGCTCGA
GGCGCTGCGCCTGCTCGGGCAGGTGGAGGCTTCACGCCGGGCCCGCGCCCAGGGACGACCCCTTACCCCGCAG
GTCCCAGCGGGACTCGGGGCCCCCGGATCCAGCGTCTAGCCACCTGTGCCCGCACCGCCCGCGAGGCCTTGTGA
CACCTACCACCCTGGCCGCCCCCGCGTCCCCCCGCGCACGAATGTAGGGATCCTGACACCCCGGAACCTAAGAC
GGGGCCCCCATACACTTTCGTACAGCGATTCGGGATTTCTCTCGAACTCTGCACATCTGTATGGCAAAGTTGA
TGGCCTGCATTATTTTTCTGATAATTCAGCGAAAGATGGCGACCAGAGCTATGCGCGTCTGGGTTTTAAAGGC
GAAACCCAAATTAACGATCTGGTCAACGAACAGAT
ACAGCATACGTTTTT

Contig 10 (3753 bp)

AG**ATTCCAATGGGGATCCCGATGAGGAAGCCGCTGCTCGTGCTGCTCGTCTTCTTGGCCTTGGCCTCGTGCTG**
CTATGCTGCTTACCGCCCCAGTGAGACTCTGTGCGGCGGGGAGCTGGTGGACACCCTCCAGTTTGTCTGCGGG
GACCGCGGCTTCTACTTCAGTAAGTAGCTCAGCGGGGCACGGGGCCGGGGCGGACACAGCAGGTGCTCCATCG
GTGCTGCCCCGGTACCTGTGCCGGTCCTTCGGGATGGATGGTGTGGGGACGGGGGGCGGGGGGCGGCCAAGG
GAGGACCTCTCCTCCGAGGGTCTGAGACTTCAGAGCGGGGGCGCCCTGGCCCTGCGCAGTGATTGGCACCTGC
CATGTGCCTGGCTGGGGCTCACACCCCCTGACGTTCCTGCAGCGTGACTCGAAACGGGAAACCGAAGGGACGG
GTGGCACGGGGTGGGGAGGCAGACCGTGAGTGGCAGGCGTGCGAGGGGTTCTTTCGGGCGGGGTGGCCCAGGC
AGGCCCCACAGGATGACAGCCTGTCCCCTCCTGCTCCTCCTTGACCTGCCCACAGCCAGGGCTGCAGGCACTG
ACATTCACCCATGGTATTGTGGTGCCTGACGTCTTGGCAGTGGGCATGGGTTCATGCACTGTTGGATTGAAAG
TGGAATAAGATGGGTTGAAAACCAATAAGAATAAAGGCGCGTGTGGCTGGCGGCATCTGCGAGAGGTGACCGC
TGCCCTCCCTGGGGTTGGGCTTTGGGTGGGTTCCCATGGGTGGGGCGGGCCGCCATGCAGGGTGCCCGCCTGC
TGGCCTCAGAGTGCTTTGCCGTCCTCATCTTTCTCTCTGGCCCCCGTCCCGCTCCTGAGGCTGGCTGGCTGGG
CCCGCGGAGACCTCCGCTCCCGCCTCGTCTGTGCCCAGGGAGCAGGGTGGACCCTCCCTTGGGCTCTTGCCTG
CACCTCCCAGCAGGCTGGGCCTCAGTGTCCTTACCTGTAGGATGGGTCAGGGGCGTCCTGGAGAGAGTCCTCG
GGACAATGGGGAGGCTGGGGGCAGGCCCAGCCTGACCCTGAAGGTGGGAGTGTGTGCTCCCCCTGGGCTCAGC
CAGCCGCGCTTGGGGCCGGGAGGGGGTGGGGGACGTGGCTGGGGCAAGTTGTCAAGGGCCGCGAGGCTCACCC
CCGCCCATCGCTCCCCATGTGGCAGCCTCTTCTGCAGCCTCTACTTACCCACCCTCTGAAATGGGCTGAAAAC
ACCCATCTTGGCATGCCAAAGCTTCTCTGTAAAAAGCGTTGCTGCTTCTTGATGCTTCTGAGGCCCCTGCCTG
CCCTGGCCTCTGAGCCCTCTCTCCTGCCTCGTTTGGGGGCAGGGAGTGGCACCATAGAATCTGGCGCTGGG
CCTGGGGAGCGGCCCCCTCGTGCCAGGCTTCCCCGAAAGGAGGGCTGGGCTGAGCTCCCGACCCTCTGGACCC
CTTACCAGGACCCCTTACCAGGGGCTTCCCCCCCCCCCCCCCCGGTGGCGGCGGGCTGGGCTGGGGCCTTTT
CCTTGCAGCCGAGTCCGAGCTGTCGGAGGCGAGGGCGAGGACGGGAAGAGAGGAGGGCGTGGTTTCTGCTGGT
CCTCACTCCTCTCCTCCCGTCTTCCTCCTCCTCCTCCCATTCCCACCTGTGTCTCCGGGTCCCGGGGCCGCAG
GCTGCCCAGGCGCCTGCTGATCCATTGGGGACCGCACTCGGGTCCCCGCTGGCCTTCGGGTCAGGGCCACGGC
CCACCTATTTTCCAAACAGCCTTGGGTCGAGGCCCAAGAGGCTGGGCCCGGTTTAAGGACGGGGAGGGAGGCG
CCAAGAGGCCAGGGGCTGGTCCCGAGCACGCCCGCACCCGCTCACCCCCGCTGTCCCCTCTCCTTCCCCGGGG
GGCCCCTGTGCACCCCACTCTCACTTCTTCTGCTCGAGGCCACGAGGCTGGCTGTCCCCGCAAGGTGACCGGG
CGTCCTGTCTGGAGGGCGGGGGCCGGGGCGGCTGGGGGCACCGTCCGTGCCCGGGGCCCCTGTGCTGACGTGC
CCTCCCCTTGGTCCTGTGGGACTTCCAG**GCAGGCCGGCAAGCCGCGTGAACCGCCGCAGCCGTGGCATCGTGG**
AAGAGTGCTGCTTCCGTAGCTGCGACCTGGCCCTGCTGGAGACCTACTGCGCCACCCCCGCCAAGTCCGAGAG
GGACGTGTCGACCCCTCCGACCGTGCTTCCGGTAAGGCAGCCCCTCTCTCGGCAGCGCCCCCCCCGGGGGG
GGCTGTCTCCTCTGAGCCGGGGGACCGGGGCGCAGCCGGCTCTTGGGCTTCAAGTGCTGCCAGAGGGGCCTTC
CCCGCTGGGGACCCTGGCCAGAAGCCAGGGCAGTCTTCGCTCTGTCGCAGGGCAGGCAGGCAGGAG**GACCCCG**
CAGAGGTTGTTGTTCTGGACAGGGGCTGGGGGCCAGGCCCCCCCTGACGGCCCTTCCCCTCTCAGGACA
ACTTCCCCAGATACCCCGTGGGCAAGTTCTTCCGCTATGACACCTGGAAGCAGTCCGCCCAACGCCTGCGCAG
GGGCCTGCCGGCCCTCCTGCGCGCCCGCCGGGGTCGCACGCTCGCCAAGGAGCTGGAGGCGGTCAGAGAGGCC
AAGCGTCACCGACCCCTGACCGCCCGTCCCACCCGAGACCCCGCCGCCCACGGGGGCGCCTCTCCCGAGGCGT
CCGGCCATCGGAAGTGAGCCAAATTGTCGTAATTCTGCGGTGCCACCATCCACCTCGTGACCTCCTCTCGACC
GGGACCGCTTCCATCAGGTCCCCCTTCTGAGATCTCTGTACCCTTCTGTCTGCGGGCATCTCCGCCCCGGGCC
CCGTGCCCCAACCTCCCCATGTCAGGCTAGTCTCTCCTCGGCCCCTTCCATCGGGCCGAGGGCATCCAAACCA
CAAACCCAATTGGCTTGGTCTGTATCTCCCCCCAAATTATGCCCCCAATTATCCCCAAGTTACATACCAAAAA
TTGAACCCCTCAACCACACCCACATACAATCAGCCCCCGTAAAACGAATTGGCATCTTTAAAACACCAGAAAA
GCGAATTAGCTTTAAAAAAAAAATAAACCCAAAATATCAATTAGCTGAAAAAAAAA:TACTAAAAATAAATTG
GCTTAAAAACAATTGGCAAAATAAAAGAATTTGGCCCCCCCCTTCCTTCTCTTTCTTTCGGACCTTGAGTTA
AATTGGCTGTGACCCATCATCCAAGAGAAAGGAAGGACCAAAATTTGCAGGTAGGCTTGTCGCCGCTCACAG
CCATCTCCCTCCTCCTGCCACACCCTCGCCGGCCACTGGCGGTGTGGCACCAAGGACCCAGTCCCGTCCTCTC
TCTAGTCCCATGACCGAGACCGCGGTGGAGTTGGCTGGGAGACCCCGTGAGATCAGAGGAGGGGAGCACGGAA
CCAGAAACCCAAACCTGCACAGGTACAACATGACTGGCCCCCCGCACAGCCCAAGACCTCTCATCTCAGTCTC
CACTTAAAAAGCACCTGTACCCACACGCATCCCTGCAGAAACACACACACACACACACACACACACGCACGCA
CGCACACACGCGCGCACGCACGCGCACACACACACTCATGCGTATACACACACACACACGCACGCACGCGCAC

FIGURE 6, CONTD.

CCACACACACACATGCATTCACACACACACACTCGTGCATACACACGTGCGCGCGCACACACACACACA
CACACTCTCTCTCTCTGTGGGATCCCTGAG

Contig 19 (500 bp)

TGGCTCTGGCATAGGCTGGCAGCTGCAGCTCTGACTGGACCCCTTGCCTG
GGAACCTCCATATGCCCGTGGAAGCGGCCCTAGAAAAGGCGAAAAAAAAAA
AAAAAAAAAAACAACCAAACAAACAACAAAGCCAAAACACACAGAACTC
ACAGACACAAGAAGAGACTGGTGGTTGCCAAAGGTGGGGTCGAGGGTGGG
AAAAATGAGGAGAGGGGGCAAAACACACAAACGTGCAGCCATAAAATGGT
AAAGTCCCGGGGACCTCCGGTAGCGCGTGTGGGGACTCGGGTTGAGAACA
CACCGTGATGTGTATTCGCGAGTTGCTAAGAGTCCCTGTTGGAGAAACAA
ATGCGTATCGACGTGTGGAAATGAAAGTTAACCCGACCTGCTGTCGTGAT
CACTTTGCAACACATACAGACATAGAATCATTATGTTTTACCCCTGGAGC
TGACAGCGTTATACGTCCCCCAGCCTCAATTTAAAAACAGCGTTGCCGTG

Contig 20 (400 bp)

TTCATACTGTGCAATGCCAGCCTTAAATGCACAGAGGAGAGCATTAACTT
CTTTGCAGAATCACTGAAATGATACCACTCATGTTTTGCAACTTGCACTT
GGGCGTTATTTTATTGGTGCCGGAACAGCGGCGATGTGGCACCAAACTAG
CGCCGCTGTTTTTATTTCCCCTCGGTATCCGCGCTCTCGCTGTCTTCCCC
CCCTTCCGCTTGCAGCTGAGGAAAGGGCTGAGAGGAGGAAAGTCTGCATT
CACCCATCTCCCCCTGCCTCTGTTGTCATCCTTCACAGAAGTGGTGGCCT
GTGCGGGGAAGTCACTAAACCTAGGCAGGTGTCCCGTGGGGTCATGCTTG
TTACACCTTTGTGCACCTGGCCCAAGTTCTGGGTGGAGCGAGAACGTGGC

Contig 21 (559 bp)

AGCTAGCCCCCCCAGCCAGGGCCAGGCCTCTCCTGCCACCCGCCCAGCCA
GCATGTCTCAAGAGGAGGGGGCCTCTAAGGGATGAGGACCTGCTCCAGTC
GGAGACACGAAGCCCCGCCGGCTCCTCCCCGAAAGTCCAGCTGCGGCTTT
CGAGCACGGCTGCGCCCTTCGTCAATCATTTCAGCCACAGAAGTGAAAGG
CGCTTTCGTGGCCGAGGCAGGCGGGACACAGAATGGAATCCCACCCCAGA
GCGAAGAGCCGCCGTGGGTGAAGCGCGTCTCTGGTGGGGACCGGGCCGGG
AACTTCACATGGGGGTCGCTGTCCCCATCTCCCCATCGTCATTACTGCAG
GGGCTCGGCCACACCCGGAGCTGCGGGGGCCAGTGCTGGACACTGGACCT
GGCCTCCGTCCTATGATGTCATGGGGGCGGGGCCAGCACAGGGCAGTGGC
CACACCTCGGGCCTCCCAGCACCAGCCAGGATGGCAGAGGGCCCCACCCC
ACCACGGGGCATGTACATCCCAGAGGACCAGCTGAGCAAGGCTTGATANG
GGCTTCAAC

Contig 22 (450 bp)

CGTGCAGGGACCCGTGCGGGCCTTCCTGTGGCCACAGAGAACAAACACAC
CATTATCTTCAGCCCCACCGCGCGGCCTGTTAATGGGTAAACTGGGGCAA
GGGGGCCCCTGCCTGAGGCCGGGGTGGGGAGCGCAAGGCATGGCCTGTGT
GCCCCAGCCCAGTCCTTCAGGGCGCTGCTGTCCTGCACCGGGGGCCCCAG
GAAGCAGAGCACCCAGCTTCTCCCCTATTCTAGAACCAGCCCCCAGAACC
CTGGACCCAGACCCAGGCCCAGGGGATACTGACAGAGCCACGGCAAGGCG
GCCACTCCACACCCCACAGAGGGGCCAGCAAACCCCAGTCACTGCGCAGC
CCATGCCCAGGGGGCAGATGGGACACGAGAGCAGCCCTCATCCACAGCAG
GCAGGGGAGTGAACTGGTGCAAAACGGGGCGGTTCCACGAAAGTTAAGCA

Contig 23 (535 bp)

TGCCAGAGACCTCAGAGCTGGGCTCTGCCTTCCCGGGCTGACACGGAGGG
CTGTGGCTTCCACCACCCCAGGCCACAGCCAGCCTGCCCAAGTCCCTGAA
GTGTCCCCAGAGGTGGCCCTGCCTCCACGCCCAACATCAGGCCTGCTGCA
GCCCTGGACGGCCCCCTGTCCCCCGGAAGCCCTCGGGGCTCTCTCGCGTC
GCCTCTGGGGAACCCTCGGTAATGTGGCCCAGCCGTGCAGTGGCCGGATC
ATTTGCTCAGGGGGGCCCAAGGCAGGGGGGTGACACATCCGCAAGTACCG
CATATGCACAGGATATGGATTGGGTGTGGATTTAACCTTTTCGCAAATGT
CTCTGCCGGTACAAATATTGTTTCTAATCCTCTGCCTCCCTGAGCCGGTG
AGTCTGCCCGGGAGCTGCGGGGAGCTGGCTTGCTGAACCTGCCCTGGCCC
CCACCCCCAAGGGAGCCCCCGGCCAGTGCTGAGGGCAGGAAGCTTGGGCA
CAGGCTGCAGAGGCCAGCGCTGGCCTCAGTCACCT

Contig 24 (868 bp)

TATTGAAGACCCTATCATGAGTTCCCAGAGCGGAGGGGTGGAAGCAGGGG
CCTACAGCCCACTCCCCATCACTCCAGACCCGTCCGGGGCTGGTGTCCCC
TGCCCCCTACTCCTGTCTCTGGTGGGCGGACGCTCGAAGGAGGCACTCTG
GCCTGGAGCCTGGAGGGTCCCTGAACTCCCGCTGCCACCTGGGCCCTCGG
GCTCCTCCTGCGCTGGGACCCGCGGTGGTGGGAAGCAGCCCTCCTCAGTG
GGAGGAGGCAGGGCTGTGGCCGCCCCGCACGGCCCTGGGGGGGACGCACG

FIGURE 6, CONTD.

CAGGACGCANGTGGGCGTGTGTGAGTCCGTCTACACGTCCAGCCAAGGGC
GGCCGCGACCGGCCAGGGTGGGCAGCCCCAGCCTCAGCAGGGCGCTCTCT
GGGGCTCAGGCTGCGCCGACGGGAGATGAGGGGTGAGGCGCAGTCTGGGG
CTGCTGCCGCAGAACCTCGCCCAGCTGGCAGCTGGGCACAGGGAGACCTG
TACTCCCAGAACCTGAGGCTGGACGTCCGAGACCCGCGTGCCGGCCTCTT
GGGTGCCTGGTCAGGGTCCTCTTTCTGGTTTGTGGGCAGAACCTCCTCAG
CGCGTCCTTGCATGGGGTGCTAATCACGGAGTAAGGAGCCAGAGAATGAG
GCACGGAGTATCCAGTGTTAACCCTGGAGTATGGAGACGGGAGTACTAAT
TGTGGAGCATGGCTCTAAGGAATGGAGTATTCGTCACGGAGAACGCGGGG
CCGGGTGAAATACGGAGAGCGGCGTACGGACAACGGGGACGGGGTATCCG
AAGGGGAGGATGGAGTATCGGCCGGAGGGTGGAGAATGGACACTAGAGGA
TGTATANNGGGCGTCAAT

Contig 25 (500 bp)

ACCAGTTTCGATGAGCAATCCCAGCGGCGTAACATTATGGCTGCAGCCTG
GTCAATGCCGGTGGAGTTTGAACCTCCACGCGTGGCGATTGTGGTAGATA
AATCGACATGGACCAGGGAGTTGATTGAACATAACGGTAAATTTGGCATC
GTTATCCCGGGCGTTGCAGCAACTAACTGGACGTGGGCGGTGGGAAGTGT
GTCGGGGCGTGATGAAGATAAATTTAATTGCTATGGCATTCCGGTTGTGA
GAGGCCCGGTATTTGGTTTGCCTCTGGTCGAGGAAAAATGTCTGGCGTGG
ATGGAGTGTCGATTGCTACCTGCGACTTCTGCGCAAGAAGAATACGACAC
GCTGTTTGGCGAAGTAGTATCAGCAGCGGCAGACGCACGGGTATTTGTCG
AAGGCCGCTGGCAGTTTGATGATGATAAGCTCAATACGTTGCATCATTTA
GGTGCTCGGACGTTTGTTACCAGCGGCAAGCGTGTTACGGCGGGTTAAGC

Contig 26 (900 bp)

ATGTTTGATGTCCGCGCGTGCTGTAAAAATTTACGCTGCTCGCGTTCTTT
GGCTTCGTCCACCACCGGAAAACGGACAAAAATTTCCGTCATACCTTTTT
CTTTCAGGCGGAAGCCAATGTCGTAATCTTCAGTAAGACTCTGCACGTCG
AAAGCAATACCGTCACCGTCAGCTAACAGTGCGGTCACGGCGCGGCGGCT
GAAACAGGTGCCGACGCCTGCGCTGGGCACTTGTCCGGCGAGGGCTTCAC
GCACCGGAACATCTTTGCCATGCAGCTCTGAAAACTCATCAATGTAAGTC
ATGCTGGTGAAGTGCGTCCATTCGCGTTCGAACGGATACACCGGGATCTG
AATCAGATCTTTACGCTCGACCAGATAGTTGAACAGACGCAATTCCATCG
GTGAAATCACATCTTCGGCGTCATGCAGAATAAAACCAGCAAAAGCGAAA
TTCGCGCTACGCTCAAATTGGGTGATGGCGTCCAGCACGTTGTTCAGACA
GTCGGCTTTGCTGGTGGGGCCAGGACGCGCGCAGACTACCTTATGCACAT
TCGGGAAGCGAGCGCACACTTCGTCAACATCACGCTGAGTATCGGGGTCG
TTGGGGTAGGTGCCAACAAAGATATGATAGTTTTCGTAGTCGAGCGTGGT
CGCCGCCAGCTCGGCCATATTGCCGATGACGCCCGTTTCATTCCACGCCG
GAACCATAATCGCTAACCGTTTTTCATCTGGTTTATACAGTTCGCGGTAA
CTCATTCGCGGGTAGCGGCGATAAACACTCAACTTGCGTTTAATGCGGCG
TACCCAGTATACGACATCTATAAAAAAATCGTCCAGCCCGCTGATGAACA
TGATGACCGCTAACGTTATCGCGATTACTTTTAAGCCGTATAGCCAGGTA

Contig 27 (500 bp)

AGCTGGATGCCCCCAGCTGTGGTCCCTTCCCTTCCCTCAGGGCAGGTTCT
GTCCCTCTTGCAGCCACCGTCACTGCTGTGGACAGGTCTGCACACCCGCC
GTCCACCAAGAGCGTGGCAGGTCCCTGGGCACGGGCCGGCTCCTGACGCA
CCATGTGTTCAAGGCAAGAGCACTGGACAGAGGGTCCAGACGTCCCCTTG
TCCTGCTCAGGCCTGGGCGGGGGCAGCCCTGGCGGGAGAGGCCCTGGGCA
TCAGAGCCTCTGTGGCCTGGAGCTTGGCGCCCTGCCCTCCCCACCTCCGT
CCTGCTCCTCGCCGCGCTGCACGGACCTCTCCCGGCCCCCCAGGCTCATT
ACTCTTAAGGACCCTAGCCCCCTATGCTGAAATGCTGTACCTCGTGCTTG
TTTTCATCTGTTTATTACCTTATCTTCATTCCTGCTTGATGATATCTGGT
TATTCTTTATTGATTATATATATCTTGTTCGTGTTTTTATAGGACACTGT

Contig 28 (450 bp)

AGTGCGGTCGGGCCGTCCTGACGCTCAACACCGTATTTCCACGCGACCGC
GGATTCAACCTGGTCACACGGACGCCATGTAGACATGTTCGGGGTTACGC
GCAGAGAAGCGACCTGCTCAACCGGCTGGTGAGTCGGGCCGTCTTCGCCC
AGACCGATGGAGTCGTGGGTGTAAACCATCACCTGACGCTGTTTCATCAG
CGCAGCCATACGTACGGCGTTACGTGCGTATTCCACGAACATCAGGAAGG
TGGAGGTGTACGGCAGGAAGCCACCGTGCAGGGAGATACCGTTAGCAATC
GCGGTCATACCGAACTCGCGAACACCGTAGTGGATGTAGTTACCCGCAGC
ATCTTCGTTGATTGCTTTAGAACCAGACCACAGGGTCAGGTTAGACGGCG
CCGGGTCAGCAGAACCGCCGAGGAATTCCCGCAACAGCCGGACGAACGCT

Contig 29 (450 bp)

FIGURE 6, CONTD.

TCAGGCCAATCTGTCTGGTCTCCAATGGGGACAATTTGGTTCTTTAGGCT
TCTGTCCAATGGTCCGAATGGCCCACTCCCCGGGCGCCGGCCAAGGGTCC
TCTGTGCCTCGGGTGGGCTGGCACGGACCGCCCCCAGGGTCGTGCCAGCC
CCGTCACCGGGGCCCAGAAGCTTCGGGCCTCTAGCTGGCTAGTCGGGCTG
CTGTGCAGGGGGGCTGCGCTGGGGGCAGAGGCGGGGGTGAGGTAAACCTC
CCAGCCGCCCGGGGTCCCTGCCGCAGCCCTAGGCGCCGAGACGGTGGCTG
GGTCGGTACCGCCAGACCCGAGGGCCTCGGGGCCCGGGTGACCCCAGCTG
TCGCACACGCTCGCAGCTCTCTTGCTCATCAGGGCTCATCCCTCTGGACC
TCTCCTACTGCCCCACCTCACCCCGCCTGGACCCCATGAAGCCCCGCGGA

Contig 30 (600 bp)

TAAAACTAGCTCTAGTAGAAACATTTTATTTAAAAATAAAAAACCTGACT
ACGTCGGGAGTTCCCGTTGTGGCTCAGTGGTTGACGAATCCGATGAGGAA
CCATGAGGTTGCGAGTTCGATCCCTGGCCTCGCTCCGTGGGTTGAGGATC
CGGCGTTGCCGTGCGCTGTGGTGTAGGTTGCAGATGAGGCTCGGATCCTG
CGTGGCTGTGGCTCGGGTGTAGGCCGGCGGCTACAGCTCTGATGAGACCC
CTAGCCTGGGAACCTCCACATGCCCTGGGAGTGGCCCTAGAAAAAGGGCA
AAAGACAAAAAAACAAAAGAAAAAGGAAAATAAAATAAAAAAGACTATGT
AAATGAAATTAACGACTGCCTAGGGTGGGATTTACAGCATGGGAAGTACA
GCATGGCCGTGACAGTGCAAGGGTGAGGCGGGAAAATGGAAATAGGTTAG
GTGAGTTTCTCCTGCTATTTGTGATGTGGTCTGCTATCGCTTGAAGACGG
ACTGCAGTGAGATAAATATGTACAGTAAGCATCCGAAAAACCGCCAGAAC
GGCAAAACGAATGACTCCAAGTAAGAACCCAAAAGAGAAAAGGAAATAAT

Contig 31 (450 bp)

GCGCGGGCGTTCCGGCTGGGGTATTTAACGTGGTCACCGGTTCGGCGGGC
GCGGTCGGTAACGAACTGACCAGTAACCCGCTGGTGCGCAAACTGTCGTT
TACCGGTTCGACCGAAATTGGCCGCCAGTTAATGGAACAGTGCGCGAAAG
ACATCAAGAAAGTGTCGCTGGAGCTGGGCGGTAACGCGCCGTTTATCGTC
TTTGACGATGCCGACCTCGACAAAGCCGTGGAAGGCGCGCTGGCCTCGAA
ATTCCGCAACGCCGGGCAAACCTGCGTCTGCGCCAACCGCCTGTATGTGC
AGGACGGCGTGTATGACCGTTTTGCCGAAAAATTGCAGCAGGCAATGAGC
AAACTGCACATCGGCGACGGGCTGGATAACGGCGTCACCATCGGCCCGCT
GATCGATGAAAAATCGGTATCAAAAGTGGAAGAGCATATTGCCGATGCGC

Contig 32 (450 bp)

GGTGGATGCTGGCGATAGCGTCATCCTCGCTTATGCCGTGCAGCGGGCAA
GGATAAAGCGCGCGATAAACATGACCCGGCATCAGCCCCATGCCCGCAGA
GTACGGATTCACCTTGCCGGTCAGCGCCAGCGTGTAATGCGTGCGCCCGT
GATACGCGCCGCTAAAAGCGATGGTGCCGCTACGTTTGGTGGCGGCGCGG
GCGATTTTTACCGCGTTTTCCACCGCTTCGGAACCGGTCGTAACCAGCAG
CGTTTTCTTGGCGAAATCGCCCCGGCACCTTCTGATTCATAATCTCGCACA
GCTCCAGATACGGCTCGTAAGCCAGCACCTGGAAGCAGGTGTGCGACAGT
TTTTTCAACTGCGCTTCCACCGCGGCCACCACCTTCGGATGCAAGTGCCC
GGTATTGAGCACCGTAATCCCGCCCGCGAAATCAAGATACTCACGGCCTT

Contig 33 (500 bp)

ACGTGAGGTTTGGGGGAGGAAAGCGGGGGACGAGCAGCCCGAGAGGAGTG
GGGGCTGGCCTGTGGCTGATCAAACTCTGAGAAGGTTAAGAGCCCCCATT
TTTGTCTTCCTCTTTTTTATTATGGAAAATTCCAAATGGATGCAAAAGTC
CCAAACCTAACTGGACATCTTCTTGGTACCAGGAACGGTCAGGCACTTAT
GATGCACCGAGCCCCGAGGGAAAAACCCTGCCGTCCTGGAGCCCACGGTC
CAGCAGGGCACACAGGCCCCAGCCCGCAAGCGGCACGGCTGAGTCAGTGA
ATGGCGTGCCCTCTGGTCAAGGACGGGCACTCTGGACCCCAGGGAAGCCT
CTGAGGAGCCCCCTTCACAGCGTCAAAAACTGTTAACAGGGCCATGTTCG
CACCCCCCCACACACGTGGTTCAGAAGCAGACCCCAGGCATCGTAATATG
TCATCCGTGAGTTCCCTGTGTGCCACCAACAGAAAGCCCATCGTCACGTT

Contig 34 (400 bp)

CGGCATCGATGTACATGGTACGCAAGGCACTCGTAAGGCCCCGAGCCTCT
AGGCCTTGTCATTGTCACGTGCTGCTCGCGGGGATCAGCAGCCAGGCTTG
TGACCCCGGCCACTTTGACAGATAAGGACACAGAGAGGCCACAGCACTGG
TCTGAGGCCCCACAGCCAGCAGCCCAGGGCAGGGAGGACTGGGTCTCACC
TGCCTCAGCTGGGCCCAGCCTCCCTGGGAGTCCCGGAGTCTCCCCAGCTT
AGGAGTGTCCCTGGAACCCTCTTCTCTCCCCTTCCCGCCCTCACCCGGAC
CCCCTGCCTCCCCCCCACCAACCCCCTCCCCCTCCTTCTTTCACCTTGAG
CTCCCCTCTGAGGACCTCTACTGTTCCTGCTTATCCTCCCCTTTGAGCCA

Contig 35 (500 bp)

TGGCGGTGAACTATGTCGTGCGTGAAGAGCATTTGTGGTCGGTAGCGCGT

FIGURE 6, CONTD.

TATATGCGGGAAGTTTAGGCGAACTGGACAGCCTGGGTTTATCCGGTAGC
GAAATCCGCTTTCACGGTAAAACGCTGCTAGCGCTGGTGGAAAAAGCGCA
GACATTGCCGGAAGATGCCTTACCGCAGCCGATGCTTAACCTGATGGACA
TGCCGGGTTATCGTAAAGCGTTTAAAGCGATTAAGTCGCTGATTACTGAC
GTGAGCGAAACGCATAAGATCAGCGCCGAATTGCTGGCATCGCGTCGGCA
AATCAACCAACTGCTGAACTGGCACTGGAAACTGAAACCGCAGAACAATT
TGCCGGAGCTGATTTCCGAGCTGGCGTGGTGAGCTGATGGCGGAAGCATT
ACACAATTTATTGCAGGAATATCCGCAGTAAAATCTTCCGAAGCCGGACT
GGGCGCGCTCAGCGCCACATCCGGCTTCGGCAAACTACAAATCCAACACC

Contig 36 (500 bp)

GATTTCACAAGCCTGACCCACGCGGAAATGCGCTAACAGCGTAAAGTCGT
GCGGCCAGAATTTTTTCGTCTCTTCGCTTTGCGTCAATTCAAAAGTCAGC
GCTACGCCATCAGCATCTTCATGATGTGATTTCAGCGTCCACGGCAGGTT
GCGGGCAAAACCGTGCGCAGGCAGACCTTGTTGTGCCGCCGGACCAAACC
ACGGCCAGCAAACCGGTACGCCACCGCGAATAGCGACGCCATTTTTGAAC
GGTGTGTTGTTGCTCAACCACAGAACTTCTTCTTCACCCGCAGGTTTCCA
CGAGAGAAGGTGTGCGCCCTGTAATGCAAAAGAGGCTTTTACCTGGGGAT
GATCGACCACAATGAGGTCCAGTTCATCCAGTTTACGACGGGAGAGGACA
GGGGAGATTTGTTCGATGACCGGAAGGGCAAAAATTTTCTTAATCATGAC
GCAGTCCTTTAACTTCATTTTATCAGGTAAAAAAAAGAGCGACCGAAGTC

Contig 37 (300 bp)

ACCTGATCAGGCTCTGCACTGTGTTCATCAGCGGAGCCGAGATATTTGAC
CGCCCCATGCATAACGGAAAGGCGTGGGTAAACCCCCGGGCGCGTTCCTT
TATCAAGATGACGTTCGAATATTCCGGCAGGTGCAGTTTGTTTATTCCAG
AAAGGCGTTGAGCGCGTATGAATATAATTCTCTGTGGGATTTGAAGCATCCT
TTTCCCTCCTTCGGTGAATGCGCTGAAAACGGCTTATTCCAGCCGGTTCA
GGGTACGCCTGATAATTTGCATTTTAAATACCATTTATTGGGTACTTTTT

Contig 38 (450 bp)

ATCCTTTTGGGGTCTGGCAATTACGCAATAAAGAAGGCCCCCATGCGATT
AAAGTCACCGGCCCACTGTCGTCTAATCATGGAGAAATTGTCCATCAGTG
GGGTCTCGATGGGCAGGGGATTGCTCTGCGTTCCTGGTGGGATGTTAGCG
AAAACATTGCCAGTGGTCATTTAGTGCAAGTGCTACCGGAATATTACCAG
CCAGCGAACGTCTGGTCCGTTTATGTTTCAAGCCTGGCGACGTCAGCGAA
AGTGCGGATAACGGTAGAGTTTTTACGCCAGTATTTTGCCGAGCACTACC
GGAATGTTTCACTGTTGCATGCCTGATTTATGATTCAATTATCGGGTTGA
TATCAGTTTAAAACCTGATTTTCTCCTTTCTAAGCCGCTACAGATTTGGT
AGCATATTCACCTTTAATCGCGCATGATCTAAAGATAATTGAAGAGGTTA

Contig 39 (450 bp)

AATGTACTGGCAAAAAGCCAATGGCGAAGCGTGGGGAACGTTACATGCTC
TGCTGGCGGATATTAATAGTCAGGGTCAGGTGCAGATGGCGATGAACGGC
GGCATCTATGATGAAAGCTATGCGCCGCTCGGTTTGTACATCGAAAACGG
TCAGCAGAAGGTGGCGTTAAATCTCGCTTCAGGTGAAGGGAATTTCTTTA
TCCGTCCTGGCGGCGTGTTTTATGTCGCGGGAGATAAAGTCGGCATCGTT
CGTCTGGATGCCTTCAAAACCAGTAAAGAGATTCAGTTTGCGGTGCAGTC
AGGGCCAATGTTGATGGAAAACGGTGTAATTAATCCGCGTATTCATCCCA
ACGTCGCCTCAAGCAAAATTCGTAACGGTGGTTGGGATTAATAAACATGG
GAACGCCGTGTTTTTGTTGAGCCAGCAGGCAACAAATTTTTATGATTTTG

Contig 40 (400 bp)

GACATTAATCATTTCAAAATCAAAGCCCCGGTTTTCCATCGCCCGTTTGG
TGGCGTGGCACTGAACGCAATCGTTACGAGTGTAAATAGTAATGCGCATG
ATTCGTATTTCCGTTTAAAATGAAGATACGGCGCGATGATACGCGTCGGG
TTGTCTCTCTGTTGATACAGAGATACTAGATGTAGTTGAAAAAAGATTCA
ACCACACAATATATAGCCCAGTAGGGGTCGAAATTACCCTGGATATGAGC
GTGACGGGGTAGGGGGATTTTTGTGATTCACCAGGCAAAAAGAAACCCCG
AAGACAGGCTTCGGGGTCAAAGACGCGTATTTATTATCATTTTTGCACTA
CGATTTGCGCATGCTTAACAGTGCGCCGATTAAAATATCTACCGCAGCTG

Contig 41 (500 bp)

GCAAAATCACGTCCGCGACCTGGCGTTGTCGCCTGGGCCATATTGGCAAAG
GAGCTGGATTGCGGTGCCTGCAAAGTGCCCTGAATAATGCCATTGTCCTG
TACCGGGAAGAAACCTTTCGGAATGAACACCCACAGCAGCACGCTAAGCA
GCAGCGTGCTGAGTGCCACGCTTAAGGTCAGCCACGGATGATTCAGCACT
TTCGCCAGTCCACGACCATAGGCGGCGATTATCCTGTCGAACATTTTTTC
CGAGGCACGGGAGAAGCGGTTCTGTTTACGCAACGACTCCTGGCTGAGCA
TCCGCGCGCACATCATCGGTGTCAGGGTCAGCGACACCACCGCTGAGATC

FIGURE 6, CONTD.

AAAATCGCTACCGCCAGGGTAATAGCAAATTCGCGGAACAGTCGCCCGAC
GATATCGCCCATAAACAGCAGTGGATCAACACCGCAATCAGTGAGAAGG
TCAGCGAGATAATGGTAAAGCCGATTTCACCTGCGCCCTTGAGCGCCGCC
Contig 42 (400 bp)
AGCTATCTACGGCAAAAGGCACGGTAGTCAATTTCGTTGTTAAATACATC
AAGCGTTTGGCGCCGAAATACCATCTGCCAGATGCCATTTCATTTCGTAG
CGCACTGCATAACGGCTACCGGATGCAGTACGTCAAACCCGAACTGGGGC
CGGAAGGATTTAGCTTTTCTGCAATACACCGGCGGCACCACTGGTGTGCC
GAAAGGCGCGATGCTGACTCACCGCAATATGCTGGCGAACCTGGAACAGG
TTAACGCGACCTATGGTCCGCTGTTGCATCCGGGCAAAGAGCTGGTGCTG
ACGGCGCTGCCGCTGTATCACATTTTTGCCCTGACCATTAACTGCCTGCT
GTTTATCGAACTGGGTGGGCAGAACCTGCTTATCACTAACCCGCGCGATA
Contig 43 (450 bp)
GATTAGCGCCAGATGCTCGCCATCGAAAAGTTGAATCAACCCCAGCTGCG
GGTAATAAGTGCGCGTACGAACAAATTCAGTATCCAGGGCTATCGCCGGA
AAGGCACGGACGGCTTCACACAAAGAAGCCAGCGCATCGTCCGTGGTAAT
CATTTGGTAATTCAAATTGTTTTCTCTTTAGTGGGCGTCAAAAAAAAACGC
CGGATTAACCGGCGTCTGACGACTCACTTAACGCTCAGGCTTTATTGTCC
ACTTTGCCGCGCGCTTCGTCACGTAATTCTCGTCGCAAAATTTTTCCGAC
GTTAGATTTCGGTAACTCATCACGAAACTCCACCAGCTTCGGTACTTTGT
ATCCCGTGAGCTGACGGCGGCAAAAAGTCACCAGTGACTCTTCGGTAAGC
GATGGATCTTTTTTCACTACGAAGATTTTCACCGCTTCACCACTGGAGCC
Contig 44 (750 bp)
GAGCAGCCCGCGTGATGACAGGCATGCGCCCCGCGTCGCCTCTCTCTCTCT
GGTGCACTGAGTCACAGGATGGCGGCGGTGGGCGCGGTGGTGGAAGCGGT
CCTGGAGGGCTCGGGAGGGAGGATGCGCTCAAGCTGGCTCCCCGTGGGGC
TGGCCCGGAGTAGCCTCCGTGAGGGCACCGTGTCTGCTCCCAGAGCCCGC
TCCCCGGCCTGCCCTGCCTCCCTTCCCTGCCCCAGTTCCCCCGGAGCCCC
TGGATCCCGATGGGAGGCGCCCCTGGGGAGACGGGACCAGGGAGGGGCCC
AGAGCTCTGAGGCCACCAGACCTGGCCAGGACCCTTCGTGGGAAGAAGAG
GTGGGCCCCAAAGGCACCTAGAGAGAGGGAGGCTCTGCTGGCTGGGGGGC
CTTCCAGGCGGGGCTTCCAGGCAGGGCCAGTGTCCTGGGGGCTGGAGGGA
GTCCCTGGCTGCTGGGGGGCGGCAGGAGCACCTGGGGCGTCTGGGAAGAG
AGCGGGAGGAGACTGGAGCCAACTGGCGGGACAGAGGAGGGGTCCAACCC
CAGCGGTGGTGTTGGGGGTGCTGGTGGTGGAGGCCCTGAGAGGCTGTGCT
GGGGGGCAGAGCGGGTCCTGGGAGGGGAGAAGGGGTCCCCAGGGCTCATG
GGCCCTTCGCAGCAGTGGCAGTTGGGGTGGGTGGCTGTCTCTAGGGCTGT
ACCACGGTGGGTGCCTGGAGAAAGAGGTCCTACCCCTAGTCTTTGCTGCA
Contig 45 (300 bp)
TGGGGACCCCACTCCAGCCCCACTGAGTGACGCGCCCCCCTGTGGTCCCA
CCGCCAACCCTGCCTCACACCAGAGGGGCTGTGGCCACACCTTGTCCACA
GCCTGTCCCTGAGACCACGAGCCCCCGGGCTCAGCCCCCTCCTCACCCCT
GGACCGAGGAGAAGCCCCCACCTGGGCTCAGCTCTTGGAGCTAAACTTCC
AGGAAGGTTCTGGTGCCCTCGGGTCTTAGAGCATGGTGGGGAGGGGGATG
CTGGTGGGGGCGCAAGCCCTCCCCACATTTCGCACTCGACCCGGTGGGNG
Contig 46 (300 bp)
CCGGCTAGAAGCCACGAGAGCCCCCAGGCCCCGCCCGACGTCTCTCCTGC
AGGGATTCGGCAGCCCTGGGGCCACAGGGCCTGAGCAGACCTTGGGGTTC
CGGTGTGACTCCAGCCAGGGTCCCTACTGTGTAGGCACCAGGGCAGAGTC
AGCCCTGCGCACCATGGCCACAGCTGCTCCCGCCTGAGCCGGGCCCCCCGC
CCAGGCTGGGCCCCCTCAGTGCACTGTCCCAAGCCAGCTGCTCTCCCCAC
CTCCACCTTCTCCATCCAGGTCCTGCCCCACGGCCTTTGCTCAGGCCCAG
Contig 47 (500 bp)
TTGACTGGCACTAGCACGAGCTCTGTACCCGGGGATCTGGGCTCGGGAGA
AGGGAGACCCCCCACCCGGCAGGCCGAGGGCGCTGTCACACCATGACTCT
CAGCCTTCCCCACCCGACGGACAAGAGTGACCCTCTCCCAAGCCCCCACT
CACCCAGGACCGCACACCCCGTGAGTCCTGCGAGTGGGGGCGGCTCAGGG
GCCCCGAGTCCCAAAGGAGTCTGCTGGCCCTGGGGGGGAGGGGAAGCAGC
AGGGTGGTCACGGGTCTCCCTGGTTGGCAGGACCACAAGCTCAGCCCGCT
GCCTCCCAGAGGGCAGCCGGACACCAACCAGTCCGGGGACCCCACGTACC
TCAGCTGCTGCAGGTGCCCCTGCCTGTACTGGTGCCAATGGGGCCGCTGG
GTGCTCCCATGGACAGCTCGCCACTCATCCCAGCCGCCTACCCCCCTTCC
GGGTCCAGTGTCCGGCCGGCCACCCGCCTGCCCAGCCCTGGCCTCCTCTC
Contig 48 (500 bp)

FIGURE 6, CONTD.

GGGGTTGCCGCAGGCTGCTGTGTAGGTCGCAGACGCAGCTTGGATCTGGC
GTGGCTGTGGCTGTGGCTGTGGCTGTGGCATAGGTCACCCACTGCGACTC
CGATTTGACCCCCAGCCCGGCAACTCCCACATGGCACAGGTGCAGCAGGC
AAAATAAATAAATGAAATAAAAATAGGTGAAGACAGTGGATTTCATCTCT
TGGGGTTGCGGTAAGCTCTACACAATAGGGAGTTTACCATTTTACCTGTT
TCAAGTGGCACTGAGTCAGCTCACAGTCCTGAGGGCCCACAGATGCCGTC
TGCCTGGGAGATTGTTCCTCTCACCACACTGCCCCTCTGTCCCCACTAAA
TACTCACTGCCCTCCCCGTCCCAAGCGCCCCTGCCCCACCCTCTGCTTCC
TGTCTCTGAACTTGCTGGCCACCAGCGACCCGTCTGGTGACCTCACTCTTC
GGCCCCATTTGTCGCACACCCCACCTGGCCTCTCCCCGGCATGCGCAGAN
Contig 49 (600 bp)
GGGATATTTGGGGGCATATTTGGGGGGGAGATCCCCACAAGGCATTTGGG
GTTTGTGCTTTGGAATGCCCCCGGCCCCGATGGAGGGGGCCGGGGAAGAA
TCTAAGCCTTACTTGGGGAGGGTTGGGCCCCGGGGCCCCGGGCCGGAAAT
GCCCCCAAGACAGAAGGTGTACAAAATTTCTCAAAAGGGTGACCCTTAAT
GAAACGGGTCCCGGTTGGAAAGAGGTCACCAGGGTGGATTGGTGGCACCG
CAGAATTTACGACATTTTGGCTCTCTTCCAATGGCCGGACGCCTGGGGAT
AGGCGCCCCCGTGGACGGCGGGGTCTCGGGTGGGACGGGCGGTCAGGGGT
CGGTGACGCTTGGCCTCTCTGACCGCCTCCAGCTCCTTGGCGAGCGTGCG
AGCGCGGCGGGCGCGCAGGAGGGCCGCGCAGGCCCCTGCGCAGGCGTTGG
GCGGACTGCTTCCAGGTGTCATAGCGGAAGAACTTGCCCACGGGGTATCT
GGGGAAGTTGTCCTGAGAGGGGAAGGGCCCGTCAGGGGGGGGCCTGGCCC
CCCAGCCCCTGTCCCAGAACAACAACCTTTGCGGGGTCCTCCTGCCTGCC
Contig 50 (179 bp)
ATCTTCATATTCATGCAGAAGACACTCTCCTGCCTTTCTATCTTGGGGAA
AAGGACGATGTCACTTATGCAATAAAGCCCACTTGCTGGCCGGGGCTTGA
CATTATTCCTTCCTGTCTGGCTCTGCACCGTATTGAAACTGAGTTAATGG
GCAAATTTGATGAAGGTAAACTGCCCACC
Contig 51 (500 bp)
CTCGGGCTGCTTCCAGGGGGCCTTGGGGAGCCATAGAATGCTATGCAGCA
AGAGAGTGCTATGGTCAGACGACTTTGGGGGAAGGTCTGGGAGAAGAGGG
GTGACTGGCCACTGTGATAAAGAGTGGGCGCTTCCTTGAGATAACACGGT
GGGCAGCCGAGGTGGACCTGTGCAGGTGGAGAAGGCCTCCTGCCGCGGCC
AGTACGTGGCTCTGGGCTGCCGGACACGAGAAAGCCCACCTCCACGGCTG
CCTCCAGGCGGCCCTTCCTCTCTTCACACCGCCGGGCCATGCCCAGGTGC
AGGTGCCATCAGAGGGTGCTCAAGAGAAGCTCTGGGCTGGGGTTGTCCCA
GGTCCCGGAAGCCCCGTGTCCCAGGGGCCACCTGAGGAAGCGTGGGCGCA
CAGAGACTGTCCCTCGGTGCTCAGAGAGGGTCCCGTCCCCACGGCAACGA
CGCCCAAGGCGGAGGTGGTCAGAGGTCTTGGGAGGGAGGATGGCCGCGCA
Contig 52 (900 bp)
TGTGTTGCACCTGTTGCTGCCTGTCGACTCTAGAGGATCAATACTCCTTA
CATAATTAAGGAGAACAAAATGGAACTTAAAAAATTGATGGGACATATTT
CTATTATCCCCGATTACAGACAAGCCTGGAAAATGGAACATAAGTTATCG
GATATTCTACTGTTGACTATTTGTGCCGTTATTTCTGGTGCAGAAGGCTG
GGAAGATATAGAGGATTTTGGGGAAACACATCCCGATTTTTTGAAGCAAT
ATGGTGATTTTGAAAATGGTATTCCTGTTCACGACACCATTGCCAGAGTT
GTATCCTGTATCAGTCCTGCAAAATTTCACGAGTGCTTTATTAACTGGAT
GCGTGACTGCCATTCTTCAGATGATAAAGACGTCATTGCAATTGATGGAA
AAACGCTCCGGCATTCTTATGATAAGAGTCGCCGCAGGGGAGCGATTCAT
GTCATTAGTGCGTTCTCAACAATGCACAGTCTGGTCATCGGACAGATCAA
GACGGATGAGAAATCTAATGAGATTACAGCTATCCCAGAACTTCTTAACA
TGCTGGATATTAAAGGAAAAATCATCACAACTGATGCGATGGGTTGCCAG
AAAGATATTGCAGAGAAGATACAAAAACAGGGAGGTGATTATTTATTCGC
TGTAAAAGGAAACCAGGGGCGGCTAAATAAAGCCTTTGAGGAAAAATTTC
CGCTGAAAGAATTAAATAATCCAGCGCATGACAGTTACGCAATGAGTGAA
AAGAGTCACGGCAGAGAAGAAATCCGTCTTCATATTGTTTGCGATGTCCC
TGATGAACTTATTGATTTCACGTTTGAATAGAAAGGGCTGAAGAAATTAT
GCGTGGCAGTCTCCTTTCGGTCCATAATAGCAGAACAAAAGAAAGAGCTC
Contig 53 (450 bp)
CCAGCCACCAGCTGGACCCTCCCGGAGAGGGGCTGCCTCCTCTTTCCCGC
CCAGACGCCCCCCAGCAATCTGTGGCCAAGAGGGAGTGATACCGAAGATG
GCCACATGGGGCCGCCAGCCCACAGGGAACCCCAGGAAGGCGCTGGACCG
TCAGGAGTCAGGGCTGCTGTGCACCCATGTGGCCTGGGGACTTTCCACAG
CCTGGTGGAGATGGCCGGGCACACCGCTGCCTCGGGGGAACGTGCACACG

FIGURE 6, CONTD.

GGTGGTACATGTGGCCGGAGCCCAGGGCACAGGGTGAGGGGAGAAGGGAG
CATGCGGGTGCAGACTCGGAGCCCGCGCGTGAGGTGCTGGGTCCTCAGGA
CACGCTCTGGGAGTGGAGGACCCCCATCCACGCCCTCACCCAGTGTGTGC
CCGCCTGCTCCCCCGGAAACCCTCACAGACACGAGGGCACACCCAGCCCC

Contig 54 (1133 bp)

ATGGCGCTCATTAGAATTCGACCTCGGTACCTTGGGATCTTTTGACCCCT
ACCTCACGCCATCTACAACATTTACCTCCGAATGAATGAGAGACACCAAA
AGCAAATTCATAGAAGAGAAAAAAAGGTAACCTGGACTTTAAAAATGTAA
ACTTCTGCTCTTTAAAAGGCAGTGCTAATGAAGTTCAAATACAAACCACA
GACCATAAGAAAATACTTGCAAATCTTGTTCTGACAAAGACTAGTGTTCA
GAACATACGACGATCAGGGAGAGGAAAACCAGCAATCCTATAAAACTGGA
CAAAGAATTGGGGGGAAAAAAAACCCACTTGGCCAAGAAGTTGGTAAATA
AGGCCATGAAAACATGCTCAACATCATGAGTCATTAGAAAAATGCAAATT
AAAATTATAATGAGATACTACTACACAGCTATTTGAATGGATAAAAAATG
TTTTAAAAACTGATTATACCCAGGTTTGGCAAGAACATGAGAAACGAGAT
TTTCACACACGATTGGTGGAAAACAGAAAATGGTCCACCCACTTTGGAAA
AGAGCTGGGCACTTCCCTCAAAAGTTAAACATACATCCAGGACCTCACAC
AGGCTTTCCACCACAGGTGTTTATTCCAGAGACATGAAAGCGCTCATCCA
CACAAAGACTCGTAAATGAAGGTTTATAGCACCGTTTGTGGCCCGAACTG
AGAAAACCCAAATGACCTTTAACCAGAGAATATCTAAACAAAATATCCAT
TCACATTAATCACCCATAAGAAGGAACGGGCTATGGGGACGGGAACCGTA
TTGAAGAGGGTCAAAATACATACGCAGCATCAAAGAAGCCTGCCCAAAGG
ACACACACTGCAGGGTTCCATGGACTGAAACTCGAGAAGGTGAAAACTCG
CCAGCAGTGACAGAGAGCAGGTCCGAGATCAACCTGATGTGGAGGAAAGT
GAACCCTCGTGCGTTGTTGGCAGGACTATAAACTGGAGCAGCCCCTACGG
ACAACAGTAGCCCGGGCTCCTCTCCTCCATCTCCCTGGGGAGCCTGAGCC
TTGAGACGCTGGGGCAAGTGCACGGCATGCTGCCTCACGTGGGGCCCCGG
TGAAAACACGTGGCAGCTGGGGAAAGAATCGTA

Contig 55 (735 bp)

TACTGCCTGTCTCTATGGACTTGACTCCTCTCGGGACTTCATGCGAGGGA
TCTTACAGAATTTGTCCTTTTGCATCTGGCTTGTTTCACTGAGCATCGTG
TCCCCAAGGTCCATCCATGTTGCAGCCTGTGTCAGGATTTCCTTCCTTTT
CAAGGCTGAATAGTACTCCACTCTGCGGATGGACCACGTTTGATTATCC
ATACTAGTAAATCCATACTAATAACTTGTTCACTGAAGCCCACAGCTTAT
GCTACCTTCCGTGGGCTCCTCCCTGCCCTGTCTCTACGCCTTCTGCTATA
GCCCCATCCCCTCTCATCCAGGCCACGCCTCCTGTCCCCTGGACACTGTC
CCAGAAGCCAACTGCCCTCTGACTGCTGCTCTCGCGTGACGCAGGACAAG
GCAGGCTCAGGGGTCCACGGGCTGGGGCCCCAGGGCTCCCCATGGCTGGT
GCCCCTTCCTGATTCCAGAAGTACAGTGGCAGCACCAGCTTTCCAGCTGC
CCCACCTTCTGTCCGCAGGCTGCTCGGGTGGGGGCAGGTGGGCAGTGATG
TCACCTGCTGTAACCACCCTACCGTCGCTCATCCCTGTCCAGGAGGTCAC
GGTGACCTTGGCAAACATTCTGAACAACACACACCTCCCTCTGCTTAGAG
GCCGGGGGCCTCCCCGGGTGACTGGGGGCACAGGCTGACCCCAGCCTGTC
TCTGTTCTCTGAAGGACATGATAAGTACTGCAACA

Contig 56 (500 bp)

AGGAAGAACAGGAAACAACGGGGTTGAGGAGAAGAAACGGGTGTCTGGCA
GGGGCACGTGCCAACGGTCCACCGGGTGCTGCCGCGCTGCGGCCTGGCGC
CAGAGGGGGCAGCTCCGCCCCTCGGGCCGCGCCCTGCCGCTTGTGCTGGC
TCGCGGCTGGGCTCTGCTTGGCTGGGTTACAGCTGGGTGCAGCCGCAGGC
TGTGGTGGGTGCCGCCGGGTCAGCCAGCCCGGCCCCACCCGGCCCCGTCTC
GCCGGCCTGGCCCGGGCAGCCCTCCTGCAGTCGAGGAGTCGCCCTGACGG
CCTGATTGGTCCACAGCCTCAGATGCAAACCAGCCCCACGTGCCTGGAGC
CAGCCAGCCCGGACACCCTGGTGGAGGCAGGAAGGCAGCAGCCTGGAGA
GCCGCGCCGGATGATGCTGCGGGGAAACCGGGCTCCCGCCGGGGGCGCCC
TGGCTCTGGCCAGGCTTGGCTTGAATGCTGACGTGAGCGGTGGCCCTATA

Contig 57 (500 bp)

TGGCGTTGCAGTGGCTCTGGCGGAGGCCGGCGGCTACAGCTCCGATTGGA
CCCCTAGGCTGGGAACCTCCATAAGCTGTGGGTGCAGCCCTAAAAAGCAA
AAAACCCCAACATATATATATATATATATATATATAATTATGGTAAAATACA
CATAAAATAGAATTTACCTTCTTAATAATTTTCAGTGCACAATTCAGTGG
CACTAAGCACATTCATGCGGCCGTGTCACCTGCTCCAGAACTTTCCATCT
ACCCAAACGGACTCTCCGCCCCATGGAACACGCCCCCTGCCCCTCCCCCG
GCCCTGCCCCGCCAGCTCCTCCCTGTGTCTGTGGATCCGGCTCCTCCAGC

FIGURE 6, CONTD.

GACCCCGTGCGTGGGCTCACAGAGTGTGTGTCCCTCTGTGACCGATCGTC
GTGTCCCCGAGGCCCGTTCTGTGGCAGCTGCGTTATGACCGACTACCTTC
GAATGCTCAGTGACTGCCGTGCATTGGACACGCAGTCCGCTACCCTTTTC
Contig 58 (550 bp)
TGCTTTCTGTGCCCCCCTCCAGCTTGGGACCCCAGCAGGGCAAGGGGTGT
ATAGGGCTTAAGGAGGCAGGGGGCGTCTCCTCCCGCTGGCTGCCCAGAGC
ACCCCCAGCCCCGCCTGCCCCTCGTCCATCTCCAGCCTGTCCTTTCCTGT
GCCCTCCCTGTCCCGGGCGGCCGCACACTGGCTTCCACCTCCCCACCCA
ACTGGCGGCCCGGTCCTTCCTGCTGAGGCACCCCGAGGTCCCCGCTGCTG
GGGACCAGCTGGCAGGTGGCTCCCACTGCTTTCTCAGCGTGGGCTTTGGA
GGGGGGATCTGCACATACCATCCCTTCAGGCCCCGTGGGGAGCCTGGGGA
CCATCCGGGACCCCTGTGGGCAGGCCCAGAGGACTGCCAGGAAGAGACCC
AGGGGACCAGGCAGCTCCCAGGCCTCTCAGCTTCAGGCCAGGGGAGCCCA
CCCCCAGGTGGCAGGTGAAGCCAGGCCCCCAACCCACAAAACTGCCCGCA
GGGAAGTAGGAGGGACAGGAGGAGGGGAGGCCAGGCCCGGGCCGCCCTTG
Contig 59 (800 bp)
TGAGGAGCGCAGGCCCAGGCCTGAGTGTGCCCAGCTTACACCCCTGGCAG
CTTCGTCCCTCCTGGCCCTAACCCCCATCCTACCCCAGCAGCAGGGGCTC
CCCCGGTGGGGCCTGGTGAGCGTCTGACTGGGGTTTGGAGTCAGGTCTGC
TCCAGGCTCAGCCCCCATCCCCAAGGGTGCCCTGCAGCACTGCTGCCCAC
CCCCTAGCGCCCCCAGACCTTCGCCCCTCCAGCCTGGATGTACCCACGGA
CCCTGAAAAGTGGGGCTGAGCAGGTGCCCTGGCTGGAGTCCCCCTGACTT
GGGGCTGGCCAGGCTGCCCTGGAGGGGCTGTGGGGGCACAGCCTGCCCCA
GGGGCCCGCTGGGCACTGGCTCTGGAGCTGACGACAGGCAGGCCCTCTCT
TCCTGGCGGGGCCACACCCTGCCCTGGGGTTTGGGGCCAAGGCGGGCACG
CCCCATGTCAGGCGGGGGCGAACCAGGTAATTACAGCCTGGCAGCCCGCT
CCCCAGACCCCCAGCCCCGGAGGGCCCCCACCCAGGCTGTGCCACCAAGA
CCTGGCATCCAGGGCCCAAAGCAGGTCAAGGGCAGCTGCTACAGATTCTT
TTAAGTTGAGACAGAATCGACACATGACAAGTTCCTGGTTTTAGGTACTT
CGCTGCCGGGGCCGCCAGTCAGTTTAGTGACCCAGCACACCCCCACACAGG
TACAATTGCTCTTCTCAAAAGAGGCCCCTGAGAGAGCGCCTGTCTTGGCT
CAGGGGTAATGAGCCCAATGGGTATCCATGAGGTTGCGGGTTCCATCCCC
GGCCTCGCCGCGTTGGTTA
Contig 60 (500 bp)
GGCTCAGGAAGCGCAGGGGCAGCGTGTGGGGCGACGGGAACCATGGGGGT
CTGTCTTCCCGCCTCTCCTCAAGCCCACCGCCCTGCTGCCCACCTCCGAC
TCTGCAGCCAGCATGCCGGCTAGAGCCCCTGTGCACCCAGCTGGTGGCCT
CTGGCTAAGGGCAGTGCTGGCTGTGGACGCGTGTCCCCTCCCCAGCAGCC
CAAGGGTCCCATCTGCCAGGCTGGTGGCTGAGGTCTGCCCTGTGTGGTCC
TTGCAAAAACCCCCGCCCTCTCCTGCCCCTTGAGGCGTGAGGGAGACGCGG
GCTGGGCGGATGCCCTCGGGCACAGCCGCCCGCGGTGGCGCCCTGTCGAG
GAGGGGGCTCCGACGTGCCCTGACGGCCCTGGCCGGGCGGAGAGGGTGAG
GCCACCTCCTGGCCACGTCCACCCAGCTGCCACGCCGCCTAGCCAGTGGC
CCGGGGCCAAGTCAGCAGAGCCAGGCTTCCGACAAGCAGAGGCTGTAGGC
Contig 61 (700 bp)
GATGAGGAAGCCGCTGCTCGTGCTGCTCGTCTTCTTGGCCTTGGCCTCGT
GCTGCTATGCTGCTTACCGCCCCAGTGAGACTCTGTGCGGCGGGGAGCTG
GTGGACACCCTCCAGTTTGTCTGCGGGGACCGCGGCTTCTACTTCAGTAA
GTAGCTCAGCGGGGCACGGGGGCGGGGCGGACACAGCAGGTGCTCCATCG
GTGCTGCCCCGGTACCTGTGCGGGTCCTTCGGGATGGATGGTGTGGGGGA
CGGGGGGCGGGGGGCGGCCAACGGAGCACCTCTCCTCCGAGGGTCTGAGA
CTTCAGACCGGGGGCGCCCTGGCCGTGCGCATTGATTGGCACCTGCCATG
TGCCTGGCTGGGGCTCACACCCCCTGACGTTCCTGCAGCGTGACTCGAAA
CGGGAAACCGAAGGGACGGGTGGCACGGGGTGGGGAGGCAGACCGTGAGT
GGCAGGCGTGCGAGGGGTTCTTTCGGGCGGGGTGGCCCAGGCAGGCCCCA
CAGGATGACAGCCTGTCCCCTCCTGCTCCTCCTTGACCTGCCCACAGCCA
GGGCTGCAGGCACTGACATTCACCCATGGTATTGTGGTGCCTTGACGTCT
TGGCAGTGGGCATTGGGTTCATGGACTGTTTGGATTGAAAAGTGGGAATA
AGATGGGGTTTGAAAAACCCAATTAAGAAATAAAAGGGCGCCCTGTGGGC
Contig 62 (300 bp)
TTTGAAAAATTTTGAGTCAGTGCAGAATTCGCATCTATTCCGCATTCAGG
CTCTCCTGTTCTCACCTTGCCTTAGTGCGGATCTTCTATAACCACCACAG
TGACGTTTTCAAGGTACTTTATTGAATAATAAGAAAAAAGTGCACACAAT
CATGTAGTTAACTTTCTGTGCTCTTTGCCAGTTTGAAGGGACCCTCTTTT

TTTCCTTTTTAGGGCTTCGCCGACGGAAGTTCCCGGGCTAGGGGTTGAGT
CAGAGCTGCAGCTGCTGGCCTACAGCACAGCTCTTGGCGGCGATGGATCC
Contig 53 (450 bp)
TCCTGGGCCACAGGCTGCAGCAGCTCACCTGGGGCCTGGGGTCTCGCTCT
GCGGATGGACCCATGAAGGCCGGAGCCAGGTGGGGGCCGAGACGGCAGGG
CAAAGGGTCTGCACACACAGCGTCCCCCCGACCCGGCTTCTCTGGGTTCT
TGGGGGGTTGGCGAGGCTTCTCTCAGTCTGGGTTTCCTGGGGAACTTTCA
AGAACTGGGAAGTCTTCCAGAAAGTTGGGGTGAGCGGACGTACCCCCAAA
GTGCTGCTCCTGTCCCCATCCCCCACCCCGCTGTCCATCGGCGAGACCCC
GGACCGCCGTCTCCCTGCCGAGGTGTGGGGTCCCCCCCTCTGCCGGCCAG
GCTGGGCAGGGGTGAGCGCCCCCTGCTCTGCACTCGGACTCAGCCTGGC
GAAGGCGGGCCCCAGGAGGTCCTGGCCTGGACGGCAGTGACCTTCCACCG
Contig 64 (500 bp)
TGTGCATCCAACCCCAGTGGCCACGGGGGGTGACCCTCGGCCGGTCAGCC
GCCCGCGTCTCCCACGGAACCGGGCCTTGGCCTGAGGCAGAAGGACCCAG
GACTCCATCCCTGCCCCGGACTCTGCCGGAGGGTGCGGTCTGCACAGAGA
CCCTGTGGGGGTGAGGCCGGTCGGGGCTGGGGTTGAGATGGATGGTCAG
GGCGGCCCCCGCGGGCCTGCAGGAGGCTGGGTGAAGGAGGGGGCCCAGCT
CAGACGCCCCCAAACCTAGCTTGGGAGAGCTGCAGCCCCGCCCCGTCAAT
CGCGACAGCCTGCCCACAGAAGGCATTCAAATGAGAGACAAATATTTGGG
CTTGAAGACTATACCCAGCCACGTCTCTTTGGGAGCCCAAGCTGCTCCCA
GGCCCTCATTTGGGTATTAATTGGTTTTCGTTTAGAGATTTGCATGCTTA
TCAATGGCCACTGGGCGGCTGGGCCTGGATGCGGTCCCAGGCTTTGTATG
Contig 65 (661 bp)
TCCCACGACCTGCCCCTCCAGGGCCACATCTGGCGACACCGTCGCAAGAG
TTGGACCGGCCTGGTGTGGCCACAGCCTCAGGCCTTGTCTGGCCGCCCAG
GCCGGCTCCAGGCTCCAAGGAGCTCCTGCCTGCCCTCCGGAACCCCAGCA
CCCCGGGCCCGCTTCCCCACCAGACCTGTTTTTCCAGGTCAAGGTCACAG
CTAATTTGGGCTTAAACTGGACAAGGAGGCCTTATCTGGAGCAGGCTCCC
GGCCCTTTGGCCTCTGCCCTGGTGGGGAGGCCTTCCCAGAGGCTGTGTGT
TGGCGCTGACCGTGCAGCCCTGAGCTTGAACCCGGATAAGGAGGGACCCC
ACCTGGGCTGGAGCCAGAGAGCCCCTCGTTCCCCAGCTCCGCAGGGTTCTC
ACAGTCCCGCCCCTGCCCTGGGGACCCTGGACGTCCCCAGCAGGTGAAAG
GTCCAGATGCCCTCTGACTAGAGGCTCCTCCGCTGTCAGACATGCTCCCT
TCCCGCACCGAGGACGAGACCTCAGCAGCCCTGCGTGGCCTGGGGTGCGG
ACCCCAAGGCGTCTCTGAGTGTGTTCTAATGGGGAGCCGTGGGGCCTCAA
CAGTGGGGGTGGCACTTGGAGGGGAGCCTCCCCACAGCTGCCCCAAGATG
GGCCCTGGACT

Contig 66 (500 bp)
TTTGTTGGATGAATGAAATCATGAGAAAGTGATTGGACCGCCCCCGTTCGT
CCAGCTGCTTGCCAGCTGCTTTGTAAAGATGACCTCTCACCTTCTCAGAG
GCCTGGCCGGCCCGAGGTGGCAGTCAGCTGAGATGCCATGCTTGTTTGGC
ACGTGGGAGGCCCCTGTCCACGGCGTGGGTGCCTCTTGTGTCTAATCAGG
GTCAGGGGGAGCAGCAGGTGCAGGGCACATGTGGGGCCGGGGCCGATGTC
TGGGGAGGGCGGGAGGAGGGGGTGTGCGGAGGCCGTTCTGGGGGTGCAGG
GGACAGACCCCAGCGAGACCCTCCCTGGCCAGGCACCAGGACAGGTGATG
GGGGGCCGCCTCCGGGGCGTGTGACAGAAGCCTCTCAGAGGAGGCCCTCC
CACGGTCTCTGGACCATCAAGGGACCGGGGGCGCTGGGCCTGGGGGTCAC
ACCCAGCTGGCCGGCCAGCCCGGGTGGGGTCGGAGGCCCGGGCAGTTCAC
Contig 67 (550 bp)
GGGCAGGAGGGGCCCGGGGCTGGTGCGGAGGGTGGAGGTGGTGCAGGAGG
GTGTGAGGCAGGGCTCACTGAGCGTGCGCGGCTGGCTGTGCCCTAGAGTG
GTTAGCACGTGCCCCCACCCTCCAGTGTCGCTCTGTTCACCTGTGCCTGG
CTCACAGGTGTGGAAACTGAGACTCGGGTGTTGCATGAGCTTCCAGGATG
AGAATCAGCAGGCTTCCCAGGCAGGGCTGTGTCCGGGGCTCTGGGCTCTT
ACCAAGGAGGGGACACCCAGGGACAGCCCTGCTTGGGGGTGTCGGGCTGG
CCAGGCTGGGTGGTCCTTCCTGTGGCTGGCAGCCCTTGGCAGTCACCCCC
TTACCCTCAACTGCCCCTCAGCTGAGACACGACCTCCCTGCAGAGCCCTG
TCCACCCAGACACTCACTCGCCTCCTCCAGGAAGCCTTCCAGGGCTGCCT
CGCCCTGGTCTCAGCAGGAGACAGAGAGAGAGGGTGGGCCCAGGAGCAGA
GGCAGGCAGCCAGAGGGGAAGCCCAGGGGCCCTCACTCACCCCTGGGGCC
Contig 68 (500 bp)
TTTGCATTCAGCTCGTACCCGGGATCCTTCCCGGGGGCTCTGGGGGTGGG

FIGURE 6, CONTD.

GGAATGGGGGTCAGAGGCAGCTGTCATCTGCCTGTCCTACCTGCTCTCAC
AGGCTGGCCCTGGAGCCCTGGCCTCCTCCTAGGGGCACATCAGGTTTTGG
GGGAGGCCCAGCCCACCGTCCCACCTCCAAGACCACAGCTGGGAGCCTGC
CCCCCAAGCCTAGACCTAGTGGGGCTCCTGCCAGCCAGGCCCCCACCTTC
ATGCTGCCACCCACCAAGGTGGGACAGTGCAGCCAGGACATCCAGCTTCT
GGAGCTGCCCGAGGCTCAGCACAGGCTGGTACCCTAGGGAGCAGGTCACC
CAGGGCCGCCTGGCGAGGCCTGCGGGGACGGGGGGTAGGGTGGGCAGCAA
AAGAACCTCTGAGCTGGGCCGGGCGGGGTCGGTGAGGGCCCCGGGGCCGCG
GGCTGTGTGCGTGGCCCCTGAGCCCGTGCAGACGCAGACCCTGGGTGGGT
Contig 69 (550 bp)
TGTGCTGCTGTGGCTGTGGTGTAGGCCGCCAGCTGCAGCTCTGATTCGGA
CTCCTAGCCTGCGAACCTCCATATGCTGCTCTAAAAAGACAAACATAAAA
TAAAATGGGTGCGCTGTTAATTTGAACACTCTGCCTCCTCCAGAGACGAG
GCCGAAACAGGCCTCTCTGAAGGTCCCACCTGGCAGGGAGGAGGAGGCCA
GCCCCGTGGGGGGCAGAGAGAAGCCCGATGTCCCCAGACACACACGCACA
GGGACCGTGGCCCCGGCTGCCAGCCCCGCGGGGGGAGGGCAAGGCCAGAG
ACTCCCAGCAGCCCACAGGACCTTGGTGGCCACAGGACACAAACACAGGT
GACGGTGGGTGAGGCCTGGCCTTTCCCCCCCTGGGCACGAGCACAGGACA
CACAAGAGCCCCAGCGTGCTGACCGCCACGCCAAGGAGCCTGGATGAAGC
TGGACACCGAGAGTCCACACTGTGTGATTAGGCTGACGTGAAGTTTAAGA
ACAAGCGGGTGGCTCAGCGCTTGAAGGCCAGAACAAGGCCGGGAGGGCAG
Contig 70 (1300 bp)
ATGTCAGGATAGTAACCTGGGGTGCTGCAGTGACAATGCCAGATCCTTAA
CCACTGTGCCACAAGGGAACTCCTTGACCTAGAATCCTATACCCACTGCA
AATATATTTCAAAAAAGGTAAAGTCCTGAGCAGAAAAGCAAAAATGGGAT
AATTCATTTCTGGAAGACCTTCCTTGTTAAAGGAAGTTTTTTGGACGTGA
TGAAGGTAGAAACTCGGAGGCACACAAAGAAAGAAAGAAAGAAAGAGCAC
TGGAAACGGAGCAAATAAAGGTAAAAATAAAGTTCATCTCTTTCTCATTT
TTTAATTGCTCCAAAAGATAGCTGACCTCTAAAGTAAAAAATAGTGGAAA
TGTAGCATATGTCTCTAGCGTAATTTAAAGTATAACTTATAGCAATGATA
GCCCAAATAAAGGAGGAATTGAGAATATACAGTTGCTGTGTTCCCATTGT
GGCTCAGCAGTAATGAACCTGGCTAATATCCATGAGGATGCAGGTTCAAT
CCCTGGCCTCACTCAGTGGGTTAAAGGATCCAGGGTTGCAGTGAGATGTG
ACGTATGTCACAGACGTGGCTCGGATCTGGCATTTCTGTGACTGTGGCTG
TGGTGTAGGCCAGCATCTGCACCTCCGATTTGACCCCTAGCCTGGGAACC
ACCATATGCTGCTGGTGTGGCCCTAACAGACACAAAATAAAATAAAAATA
AAAGAGAGAGAGAATATACCATTGTAAATTTCCTCACATGACACAAAGAG
CAATGTGATATTATTTGGTATATGGTGATTGATTCAAGATGTATATCATA
ATATTGATTCAAGATGTATATATTCCTTTTCTAAAAAGAGATTTATACA
ATAAGGCAAGAGTGAAAATAAAGTGGAATGCTAAAGAATAGTTAATCCAA
AAGAAGGCAGAAAATGGGGAAAAGACATATAACAGATGGAACAAATAAAA
AAGAGCTAATGAGATTGTAAAATTTAATCCAAACATACAGATAATCCCAT
TAAATTTAAACACTCTCAACACATTGATTAAAAGAAATTGTCAAATTGAA
TAAACAAAGCAAGACCCAACTAGATGCAGACTATGAAAAACCCACTTCAT
ATAAAGACATGGGTAGGTTTAGAGCAGAATGATGGGGAAACCATGTCACG
CAAACATTTGTCAAAATAAAGCTGGTGTGGCTGTATTCATCTCAGACACA
GCAGACTTCAGAACAAGAAACACTGCAAAGGATGAAAGAGATACTGCATA
ATGATAAAGGGATCAATTTTCCAAGTGCAGGCTCCAAACAACAGAGGTTT
Contig 71 (500 bp)
ATGACCTCATACTGAATCGAGCTCGGTATCAGGGGATCTCTCAGCTGGGG
GGGAGGGCAATGGGGCATTTGTCTGAGGATGCCCCAGGGCAGGCCCATTG
GCTGGTTTGGTGCCCATGCCCCCCCCACACCCCGGCAGTGCCCCCTGCTG
AGCCTGGGACCCCCTCTGGGAGTTAGGGATTGGGGGTGGGAACCAGGCTT
TGCAGTAATTCCAGCCCCCAGGGCCCTTCCCTCCCCGCCCTCAGGACCCC
CAGCCCCGCCCCACACAGTCTCCACTGTGACAGCCTCACCCCTTGGGTCA
AGTCCTGTCCTCTCCGGCCCCCGCTGGGCAGTGGAGCCAGCTAGGTGAGA
GGCACAGGCCACTAGGGCGGTGGGCACTGCTGAGGACAGAGGGGCCTGGG
TGGCCTTGGACGAGGCCCAGCGACGCTGAGACAGTGAGCCAGGCTCCAGG
CTTTCCCAGGGAGGGTCCCTGAATGTCCACTTCTTGTGACATCGGGTGAC
Contig 72 (550 bp)
AAGTCCATTAGGGAAGGGATTTGTGCAAACACAGAGACAGGTGCAGGGCT
GGGCCAGCTGCTGGGCTGGGGCTCCTCAAGGCGCCCGTAAACCCCTCCC
TGCCAGCCGCCTGCCGCCAAGGTCTGCTGTCCACCCCGGCCGGGCTGCTG
TGTTCCCGGCGTGTGTCCTGCGAACCCGACTCCCGTTCACCCCTGAGCAC

FIGURE 6, CONTD.

TGCCTGGAGGCCGGCTGCCCAGGCGGGACGGGCCCTCAGGGCTGGGCTGG
CTCTTGGCCTGTGTTTCATTTCTGACCAGGTCCTTCTCAGTGGGGGGGC
CTTGGGTGAAGCAGGCATGTGCACCACTGGGGCCCTGTCCCCAGTGGGCA
TCCTGGGCGCTTGTCTGGCCCCCAAACCCCCAGGCCGTGTGCATCATACC
TTCACCCTGAGCCCCAGCCGAACCCCGGACATGTGCTGGGGGACCCTGGG
CACAGGGGTGAGGGAGCAGTGGCCTTGGTGGAAGCCCAGCCTTGGCACCT
GGGGAGGGGGTGCATCTGGCATGCTCTGCTGTAACCAAGCCCAGGGCAGG

Contig 73 (950 bp)

GACGTGCAGTAGCCATGACCTCTACGGCCCCCACTGACCAGCCCGTGTCC
TTGTCCCGAGACCGACCCCTAAGCAATAGGATGCAGCAGAAGTGACAGAA
CGGCCTCCGCGATGAGGTCGCAGAGGGCTCTGGCTCTGACTCAGGCCCCT
CATCCCTCGCTCTCCTGGAGCAGGGCCAGGTAGGGGCCCCCCAGAGACGC
CCTAGAGGAGGTGACGGGCAGCCAGCCCGCCCCAGGGAAGGCCTGGGGAC
ACCAGGGAACAGAACGGCACAGGCTCCTGGCACAGTCTCCCAGGAGCCCC
CTGGTGGCACAGAAATCCTGACCGGCCCAGTGGAGGGGGCTGGGGCGGGG
CTCGGGGAGGAGGGACTGGGTGAGGCCGTCTGACTCCTGGCTGAGCGCCG
CATACTTGCTGCCTGCCCACGATGCCGGGCCAGGCCTTCCGCACGGACCC
AGGCTCACATTCGCCCTACATGCCACTGTGTGGGAGTTTGGGATGGTGTG
CCCGCTGGGCCCGGGGGTCAGGGCACGCTTCCCAGAGGAGCGGGTTCCAG
AAGGCCCAGGTGGAGAGGCGATAGGAGGGCTCCAGGGGGCTTCCCAGGCC
ACCTGCGAGGACCCTCCTGGGGGGAAGGGAGCGGAGGGAGACAGCCGGGT
CCCTTAGGCCAAGGCTGAGTTGTGACCGCAGGGAGAGGAGAGAAGGAGCA
CCCACAGCAGGGCAGGGGCTGCGGGAGGCTGTGCTGGGTGGCCGGGTGGT
GGGTCTGGGGGCCAGGACCGTGGGAGGCCTCGAGGGGGGACCAGGCACGG
GAGGGGCCCCTGGACGGCAGAGTCCCTGCTCCAGCTGCCGCCCCGACCCC
AGGTCCACCTTCATTTCACAGCCTGGCCCCCGGCCGCTCTGACCGGCCCT
GCCCATGCAGGTGTAGCGGGGCAGTGAGGGCCAGGCTCCGGCCGTCCCAA

Contig 74 (450 bp)

GCAGGCCTGGCAGCAGGGAAATGATCCAGAAAGTGCCACCTCAGCCCCCA
GCCATCTGCCACCCACCTGGAGGCCCTCAGGGGCCGGGCGCCGGGGGGCA
GGCGCTATAAAGCCGGCCGGGCCCAGCCGCCCCCAGCCCTCTGGGACCAG
CTGCGTTCCCAGGCCGCCGGCAAGCAGGTCTGTCCCCCTGGGCTCCCGTC
AGCTGGGTCTGGGCTGTCCTGCTGGGGCCAGGGCATCTCGGCAGGAGGAC
GTGGGCTCCTCTCTCGGAGCCCTTGGGGGGTGAGGCTGGTGGGGGCTGCA
GGTGCCCCTGGGCTGGCCTCAACGCCGCCCGGTCCCGCAGGTCCTCACCC
CCCGCCATGGGCCCTGTGGACGCGCCTCCTGCCCCAGGCTGGGCCCTTGC
TGGCCCCTCTGGAGCACCCCGCCCCCCGGGCCCAAAGCCTTTCATGAACA

Contig 75 (1363 bp)

CCTCCAGCTGGGCCCGGCAGGGCACCGTGCCCCTCAGGGGACACCACGGG
GGGCCACAGTGGCCTCTCCTGCTCCAGGCTCTGCTCCCGCCTGGGGCCCC
CTGGGCCGCCCGCCCATGGCCAGGGCAAACTCCCAGTGCGGCTGCCCGTC
TGGGCAAAGAGGCCGCCAGGCCCCGCGTGGTCTTAGCAGGCACTGGCGGA
TGCCGNTAACTAACCATTTCTTCCGCAGGAGTCCGAATCTGCTCTGACCA
CGGGCCCTAAAAATCGCTCCTGGCCCGCAGAGGATCCCCGAACAGCGGGG
CTGCCTCCTGCTCCTCCTGCCGGGCCGGCACTCGGCAGGCACGTGCCCTC
GTCGTCCCCAGTCTGTCAACCGTCCCGTCGTTACGATCCCCAGAGTCCCA
CGCGCGGGCAGCTCTTTCCACACCCCGCACGGCCCCGGAGCTGCCTGGGC
ACCCAGATCGCCCCTGACGCCTTTGCTCCTAATTCTGCTGAAATACACAT
AACGTCTCCTTGAACGTTTGTCCATTTTCACGGGGACAATTCTGTGGCCG
TAGGTACACTCCCCTTGGGGCGCAGCCATCGCACCATCCGCTTCCAGGAG
GTCCCGTCGTCCCAGATGGACACTGTCCCCACTGATCCCTAATTCCCTGT
CCCCCCCAGCCCTGCCCTTCCTGTCTCTGTGGCCCTGGCGCCTCCAGGGA
GCCCCTGTGCGTGGATCACAAAACGTGTGTCCCTTTGCGTCCGGTGTGT
GTCTCTGAGCATCCGGAGCTTGGGGTGCTTCCACGCTGCGCCTGTGTCAG
GACGTCCTTCCCTTTTGCGGCTGCGCGATGCTCCCCGTGGGGCTGCCCCA
CACTGCGCGTGTTCGCTCATCCATCCACTAAGGCTGAGTTACTTTTGGCG
GTTGTGAATACTGCTGTGTGAACACGGGCGTGCAAATACCTGCTGGAGGC
CATGCTCTTAGGCCTCTCGGGGGGCACACCCAGAGCGGATATGCTCAATA
AGGTAATTCTGTGTTTAGCTTTTTGGGGAACCATCAGGCTGGTCTCCAGA
GTGACGGAGCATGCGTCGCATTCACAGGAATGGTGCTCGAGGCTTTGAGG
TCTCCACCACTCGCTTCCTATTTTCTGTGCGTCACAGCCGTCGGAACGGC
TGGGTGGTGCCTCTGTGTGGCTTCAATGTGCTTTTTCTTTTCCTGGCTAT
GAGGTTGAGCGTTTTTTATGTACTTGCTGGCCATTCGCAGGGTTTTTGGG
GTTTCTTTTCTTTTTTGCCTTTGGGGACGGCGCCCAGAGCGTATAGAAGT

FIGURE 6, CONTD.

TCCCTGGCTGGGGACTGAATCAGAGCTGCAGCTGCCAGCCTAGCCCACAC
CCGCAGCAACGCA
Contig 76 (500 bp)
TCATGCCATCGCCACCGCCCCCCACCCCGACGTTTCAAACACCAGAACCA
CCCCTCGGGCGGCAGAGAGAGGACCGGAAGGAGAGACAGCCTGGTCCCAA
GGCCTCGCCCGGTCCTGTGTCTCCGAGCGACATTTCTTTCTGTTTCCCTC
CTCCGCGGTCCAAGTTTCACCCATCAGAGGCGCATTGTTTTCATCATCTG
AAAAAAAAATCTCTGTCTCTTAATAAAACACAAGAAAAAGTAGCCTTCGA
AAGAAAGCACATGAATGATATGTGCTGGCGACAGTGCTGGCGGCCTCTGA
GCCGTGGTGGGAGGTGGGAGCCAGCGGAGCCCCTGACCGATCACGTGACC
CACGTCTCTCCTGCACAGCTGGCTGCACCTGCACGCGGTGACACAGGGAC
CCAGCCTCCTGCCAGCAGGTCACCCCACCCCGTCCGTCTCCTGTGGAAGG
CGCAGCGTTGCCTTCTGAGGGTGGGCTGCTCTGAGGGGCGTCCTTTGGCC
Contig 77 (626 bp)
GCCATCGGCTGCGGCGGTTCACGCGGCTTGCCGGCCTGCCTGGAAGTCCC
ACAGGACCAAGGGGAGGGCACGTCAGCACAGGGGCCCCGGGCACGGACGG
TGCCCCCAGCCGCCCCGGCCCCCGCCCTCCAGACAGGACGCCCGGTCACC
TTGCGGGGACAGCCAGCCTCGTGGCCTCGAGCAGAAGAAGTGAGAGTGGG
GTGCACAGGGGCCCCCCGGGGAAGGAGAGGGGACAGCGGGGGTGAGCGGG
TCCGCCCCTCCTCGGGACCAGCCCCTGGCCTCTTGGCGCCTCCCTCCCCG
TCCTTAAACCGGGCCCAGCCTCTTGGCCCTCGACCCAAGGCTGTTTGGAA
AATAGGTGGACCGTGGCCCTGACCCGAAGGCCAGCGGGGACCCGAGTGCG
GTCCCCAATGGATCAGCAGGCGCCTGGGCAGCCTGCGGCCCCGGGACCCG
GAGACACAGGTGGGAATGGGAGGAGGAGGAGGAAGACGGGAGGAGAGGAG
TGAGGACCAGCAGAAACCACGCCCTCCTCTCTTCCCGTCCTCGCCCTCGC
CTCCGACAGCTCCGACTCGGCTGCAAGGAAAAGGCCCCAGCCCAGCCCGC
CGCCACCGGGGGGGGGGGGGGGGGG
Contig 78 (500 bp).
TACTCGGGTTTGTTACCACTGAGCCACAAAGGGAGCTCCTAAAAATAATA
ATTTTCTTAAAGCCAATGACATGGAGAGCAGTTAGGGTGGAGGCTGGTGG
GTGGTGGGGCCGCGGCAGGCGCCCTGAAGGTCCTGAGTGGCACCCTTGGC
CGGGGGAGGTGGGTGGGCGAGGGGTGTTGAGAAGGGGCAGGGCCTCGTGG
GGGCAGGAAGGAAGAGAGCCAGTGGCTCCCAGTCCCCTGACCTTGCTGCCTT
GAGCCTGGTTCTCCCCAAAATTCTGTCTGTGTCCCTTCACTTCACGGAAG
CTTGGGGCCCGTTGCCAGGGAGACAGATGGGCTGGTGACACCCAAAATGA
GCCACCAGGAGGGGGGCACTGACTTTAGCCAGCCGGTCACATCAAGAAGC
AAACAGGCCCCCCGCTGCTGTAAAGGCAGCTTGGGGCTGGGGTCCGGGAG
CACCCCCTGGGCTGGGGAAAGGGGGTCCTCTCAGGCCCCCGGGGAGGATG
Contig 79 (427 bp)
TCTATTCGCCGTGGCCGGAAGAGGCTAACCGTACATTGACCGGGCATCTG
GCGATGTATCACTTCTCTCCAACCGAAACTTCCCGGCAAAACTTGCTGCG
TGAAAACGTTGCGGATAGCCGAATCTTCATTACCGGTAATACAGTCATTG
ATGCACTGTTATGGGTGCGTGACCAGGTGATGAGCAGCGACAAGCTGCGT
TCAGAACTGGCGGCAAATTACCCGTTTATCGACCCCGATAAAAAGATGAT
TCTGGTGACCGGTCACAGGCGTGAGAGTTTCGGTCGTGGCTTTGAAGAAA
TCTGCCACGCGCTGGCAGACATCGCCACCACCGCACCAGGACATCCAGATT
GTCTATCCGGTGCATCTCAACCCGAACGTCAGAGAACCGGTCAATCGCAT
TCTGGGGCATGTGAAAAATGTCATTCT
Contig 80 (650 bp)
GGCGTTGCCGTGAGCTGTGGTGCGGGTCACAGATGGGGCTCAGATCCCGC
GTGGCTGTGGCTCTGGCCTAGGCCGGTGGCTGCAGCTCCGATTCGACCCC
TGGCCTGGGAGCCTCCATATGCTGCGGGAGCAGCCCTAAAAAAAAAAAAA
AAAAAAAGGAAGAAAAGAGAAGAAAGAAAAGAAAAGACAAAAGTCAAAAG
GAGCTCCCCTGAGCGATGTCTGTCTACGAGCAGGTCCCTGGGAGCCTGAG
GCAGGGTGAGCCTGGACCCCTGAGGGCCACTCCAGACTCAGTGCTCTCAC
TGGCCAAGGTCTTTGGGGACCGGCTGGGGGCGCGCGCAGGCTAAGGAGGA
GGTCAGAGGAGGGGCTTCAGGCTGCAGGGCCAGCGGCAGCTCTGGGCCCG
GGGCGGGGGGGAGATGGCCTGAGGGCCTTGCGGGGGCTGGAGGGTGGGGG
GCTTCCTGGAGTGGGAAGACGGGAAGCCAGGTCAGAGGAGAGGAGCGAGG
GCTGAAGCTCCTGGAAGGCGCTGGCTACCCCCAGCTGGCCCGCCCCCGCTG
CCACATTCAACAGCCACCCGGCCTGTGGTCCTGGCAGGGTCCTGGCAGAA
AAGCCCCAAGGGCCCCAGCCTGGCCCTCTGGGCCTAAAGAGCCAAGCCCC
Contig 81 (550 bp)
TTAACCCACGGAGCAAGGCTGGGGATCGAACCTGTAACCTCGTGGCTCCT

FIGURE 6, CONTD.

CGTCGGATTCGTTAACCACTGCGCCACGACGGGGACCCCCCAGGGCTGGC
GTTTCCCTCTGTGTGCACACAGTGGACCTGAGCCAACCAGCAGGCCCTTC
ACCACCACGGCGCAAGAGTCGGCAGCAAGAGAGCAGTGTCTCATGGCTCA
CTTTCTCCCCCTTCCCCGGAGTGGTGACAAAACCCCGCCGCCACCGGACT
CGGTTAGACAAGGCGGTGCCCAGTGCCCCCGTCTGTCACCCGCACGGCAC
GGCGCTCTCCTTTCTTTCTCGGGGCTCCACCACGTGTCCTCAGTTTCCGC
ATGAGAGTACCGCGGCTGGCGGGGTGGTGGCTCTGGGGTCGGGGGCCGTG
AGGGCAGGGCTGGGCTGGGGAGGCAGGTCTTGGCCCATTACGCGGGGGG
CAGACTCCACATCACACGCTCTCTGTGCCTCTTGGCTGCCTGACACCATG
GACTTCAAACAGGAACAGCCGTGGAGGCATTGCAGCCCAGGGCCCGGCTT
Contig 82 (550 bp)
TGACACCTCCAGGCAGGAGGGTGCAGGCTGGGCTCCCAGGTAATGGTGTG
CTGGCCTGTGGGGCGTGGGCTCAGCTCTTAGGATGGTGGGCTGGGCGCCG
ACCCAGCAAGGACAGGGTGATGGCAGGTCGTGGGCTCAGCAAATGAGTGC
CCAGGTTGTGGGGGTGGGCACTTGGGGCTCAGGGGAAGCTCATCAGCTTG
GAGAGGGACGGGGGAGGGAGGGGGCCTTGGCCAGCTGGCCCAGATGCCTG
GATGTGAGCACTCACGTGCCCCGGGGTCCACCTCCCCTCCAGTGCCATCT
GGGCAGGAGGCTCCGATGCCTGTCCCTGGGACCCGCTGTCCTGAAATGAG
GTTCACTTGGTGCCTTCCCCAGAGATGCTCGGTCCGGAAGCTGACGAGGC
AGGAGTGCACAAGGGTCTGGGGAAATGGAGCAGAGTGCGGCTGGGGCACA
GAGGCTGCCCCCAGCCTGGGAAGATGGGGAGCTTTGCAGGGGTACCCCGC
CAGCTTGTGGGGCCCTGGATACCCAAGGGTGTGAAGAGGCTGAAGAGCGA
Contig 83 (984 bp)
CTGAGCCCAGCTATGTAGATTAGACCCCGGTCCGTCCCAAATTCTTCTCA
AAGCTGTCCCGAGATGAGAGATGAGGTTTCGTGTCCTGTGCTCTCCTCG
CTTCCCCTGGGATGTGCCCTAGGGTGGGAGAGGGTGTGTCCCAGGGCTCA
GCAGGCGGTCCCATCTTCCCGAGACGGGAGAGATCCCCTCCTTCTCGGCG
CCTGTCCCCACGGCCCCCACAGACACCCCCCCCCCCGGCATGGCACCCAT
GCACCTGCCATCGTGCCCAGTAGGGGATGGGTTTGGCGAGACTCGAGATG
GCTGTAGCCAGTGAGACATGCCCTGCCACGTAGCCTGACCCCCTGGGTGT
GCTCTGTGAGATCTGGGGACCCCCAGCACACCTAGGGATCATCTTTGCCA
GCCTCCTGGGGAGCCTCTCAGAAATGGGGGCCCCCAGAAGGCTGGCAAAG
GTGATGGGGAGCGTGGGAAGTCTGGCGGTTGGCGGGGTGGGTGGGGGGCA
GTGCGGGCTGGGTGGGGGGTGCTCCGGGGTCGGAAGTGCTCCAGCAAGGT
TTTGGACACAAAGTCAGGAGGAAGGAGTGACGAGGAGACTTGCAGAATTA
CAGGTAGAATCAGGAACCCACATCGACGCCAATTGATCTATCCCCCCCTT
TGATTCTTTTCTCCTGGGGCTTTTTTCCNTTTTTTTTTTTTTTTTTTTTT
TTAATCCCTCCTTAGCTTTTTACGCGCTCAACACCAAATTAAACGTACTC
CCCACCCCACGTAACAGGGGGGCGGTGACCCGAAGGACGAGGAGCACACG
AAGCCACCATCCGTCACCTTGGCGGCACCAGCCGCTGTCCTGCCCTCCGC
CCATTTATCGCCCTTGAATTGATTTTTGTTTTGCTCTGTCCCTGTCGCTT
GGGTAGAGTGGAAAAGGGAACCTCTGTGGGGGTGCCAGCCACTGGGCCCC
CCAAAGATTTCAGGGGAATGAAACGGCTGCCGCC
Contig 84 (550 bp)
TGCCCCTGACAACCCTGCCCTGTTAGCCACACTCGCGACTAATAAGGCGA
GAGGTCAGCGGGCAGCCCCACGGGGAGAAAGTGCCTCCGTGCCCCCCACC
CCTGGCTCTGATGGCCCAGCCTGGCACCCCAAGGTGGCCTCGGCCTTCCT
ACCTCCAAGGTCCAGGCGCATGTCCAAGCACCAGCAGAAGCTTCTCCAGG
GTTGGTGCCTGCTCAGGGCAGAAAGCAGGGGTGAGGCTCCCCAAAGGGCC
ACTGGCACCAATGCCCCCAGGCAGCCCCAGCGAAGGGGACAGCCCACCCC
CAGCCCGGGGACGCAGGCCTGAGGGGACATGGGGAACCCAGAGCAGGGCC
AAGGGGAGCAGAGCCCCTCCTCCGGGACTTGAAATCTTTCCCGGGGGGCC
CAGGGAGCTGGGGTCTGCAGAGGGCACTTTCAAAATACGGCCCACCCCCA
AATTGCCACGTGGGCCACAGAGCAAGGAGTCGCTGCCAAAGTGGCCTGGC
TTCAGCGCAGGAAGTTCCCCTCCTGGGGCCTCCCCTCCTATAGGCACAGG
Contig 85 (500 bp)
TGAGCCAGGGCCTGGCCCAGCTAAGCCCCTGGAGCCCTCCCGGCCTGTTT
CCTGCCTCCCATGCTGGCGGAGCTCGGCTTACTGAGCGGGGGCCAGGCCA
GTGTGCGTGTGGAGGTAGATTCCACTCAGCTGGAGGTTGAGGTGGGCAGG
GGGCCGCAGACCCTCAGGCCAGCTCTGGCCGGCCAGGTCCCTGAAGCTCC
CCCGGCTGGCCTCCCCGTCCCTGCCTCTGGCCTTGTCCTGGCCCTTGCCT
GACAAGCTTCTGTGGCTCTGCCTGCAGGAGAGACACTGGCTCCCCCGCTC
TCGGATGAGGACGGGGCTTTTCTGCACAAGTCCTGCCCCAGAATGTTTGG
GGCGCCAGCAGCTGAGCCCAGCACGTCTCCCCCTGCCCCTGGCTGGACAC

FIGURE 6, CONTD.

GAATCCCGGCATCGAGGCGGGAAGGGGGATGGAGGGATGGGGCCTACCCA
CCCCTGCTCCCCACCCAGAATAGCTGGGCGGCCCCCATGGAGGCCGCCC
Contig 86 (913 bp)
CTGTTTTCACGTCTTCTGAGGACACACCCAGAAGAGGGGCTGCAGGCGCC
CATGGTGACTCCATGTGTTCACTGCTGAGGCCTCTGCAGACCGTCTCCCG
CAGCAGCCGCACCCGTTTCCATGCCACCAACAGCGTGCGAGGCCGCACTG
TCCCCACGGCTGTGCAACTGTTTTGAATCTGAGTTATATAAGCAACAGAC
GCTCCTTCAAACACACTCACGTGCACACGTGCGCACAGGCGCACAGACAC
ACACACGGAGTAATAGGCCTCCCCCCCCTCCCTGAGCCCAGAGGGGGCCT
GGGGCCCTGGAGCCTGTGCTTTAGGGCCTTTTAGGAAAGCTGGTGCCTCC
CAGAGGGGCCGCCCCGAGCGTTGGCTTCCCAAGTCCCCACCAACCCTCGA
CAGACTCAAACGTTGGTTTCTTTCGTGCTTTTGCCCAAGGGATGGGCCCG
AGGTGGCCCTGCCTGAGGTTTCAGCCCAGCGCCCCAGGCACCCTTTCTCT
CCCGGTCCCCGGCCACTTCATGGGACAGCGGGCCTTCCCCCACGTTGTCC
CCTGGGTTGTCGTGCTTTTCGTAATGAGACGGAGGCAGGTGCACCTGTCC
TGGGGTGAATTCTCTTCTGCAGGAACTCGCTTCCCCGGCGCCTGGTCTGT
CTGTTCCTCGGTTGTTGGAACCTCTCGTCACCAGAAAGGGTGGCTCTGAC
GTCGCCCCTTTCCCTCCGTGGCTTTGCAGTCTGGGTCTTGTCGGGGAACC
TGCCCCAAAGAGGGGAGTGACCCCCCACGAGGGAGACGTAGCTCCTGTGG
CGACAGCACCGGGGGCCCCCAGATTCATGGGGTTCACGCTCACAGTCGCA
TGACGCTGCCTTTGGACGAGGGCAGCTCAAGGGAAGCTTGTTTCCTGCCA
CGAGCCACAGGCA

Contig 87 (650 bp)
TCCACACCTGTGGAGCCGCTGCCTCGCTGATGCCCTCTGCCCAGCTGATG
GTCAGGTGCCCAGACTTGGGGCTCAGTCCAAACAGGGGCCCACAGGTGCT
GCACCTGGGCAAGGGAGCCTGTGCGCAGGGCCTCAGGTGTCCCAGGCTCG
CTGGGACCGAAGCGCACTGGGTCCTGGACTCCGGGCTTCCCCAGGGGCTG
CTCGGGGCCACCTGGAAATGAAGCCCCACCTGGCTCATAGGGTCCACGTG
AGGGCCCTGAGGCCACCAAGCCACCAAACAACTCAGTTAAGGGAGGGGAG
CTTGGGGCTGCTAAGCTCCAAGCGGGAAGCGGCCGCACTCAGCACTGCCT
CTCTGCCAGCCAGCCGCCCAGCTTGCTGACGTCCCAACCAGGCCAGGGAC
CCTGTCCCACAGATGCTGGGCCCTTCCAGTCTCTGCTCCCTGGAGGCGCT
GGGCACTGTGTGGGCACACAGCCCGCACCCGCCTGTAAGGAAGGGAAAGG
CCCCATCCTCAAAAAAGCCGTGGGCAGGTGGGCCATGATGGTCCTCCGAG
GCAGGTCCTCCTGGGACCCCTTGCTCCCTCGGGCTCGCCCAGGAGCCGCC
AGGTCTGCCCTGGATTAACTCTGCCCCGCATGTCATTTTCAAACTGGCTT
Contig 88 (700 bp)
TGGGGCCCTTTGGGGCCGGAGCGGCCAGTCTGCTGGGCCCGGGAGCAGGG
GGTCTCTGTCCGCAGGGAGGGGGCCTGGTCTCAGGGGAGGAGAGGAGGCA
GGTCTCACCTGAAAGGATCTGCCTTCTCCTCAGGCCTCTGGGATGCCTGG
GCAGAGAAACCAGAAGGAAAGGCCCAACTTGCTGGCTGGTGGGGATGGGG
CCGGGGGTCGCTCCCGGCACACCCCCCCCAAACCCCACCTTAGTGGCCAA
AGTGGGTGTCATGATGGCCACTGACCTCACGGGGGCGCAGGAGACAACAA
AATTTCAGCCACTCTTGGGGGAAGGACACTTGTGGCCTGAGTCTTAGGGG
CTGAGTTTCGGGGGGGACCCCCAGCTCTCCCCCCAGTATGAGACACCCTG
CCCACTCCTCCCAGCTGCTCCCCAAACCCAGTGCTTCTGGACGGCATCT
CCCCGCTGCCCCTGCAGCCGCTGTCCTCTGACCATGTCCCCTCCCCACCT
CCCCTCTGCAGGGCCAGGCCTCCAGGGAGCAGAGCCGAGGCCCCACCCTA
GACTGAGCTGGGGACCGAGACCCCAAGTCGCCACCCGGTCTCTGCGTTAG
AGAGGGGGTTCCGGGGGGCACCCTGGGGCGGCACTGGGGGGCGGGAAGGA
GAGCCCTGGGCCGTTCTGGGAAAGGTCTGGGAGGGAGGGAGGGGTTTTGC
Contig 89 (1400 bp)
GCACACCCGGAGAACAGAGGGAGGGGTCCTTACCAGTCTCAGGGTTTTTT
TGGGGATTTCTTTGAACTTGCCCTATTGGTTTCGAGGCTTCTGTTCTCTC
CAATCCCCCCTTCTGAACCCCCCCAAAAATGGGTTCAGCCCCCACCCCAG
CCAGAGGAAACCAATTGGGGGATTGGGGGGAGGCGGGGCCAGCAAAAGCC
TTGGGCCCCCAGCCCCCCTGGCTTTGGCCTCTGGCCTGCCAGGTAGGGGG
AGGGACGCGGTGACCTCCGGGGGCCTGGCCACGGACTCTGCCCCCACCCC
CAGGGCAGACGTGCACAGGAGGGGAGAGGCTCCGAGGAATGAGGCCATCA
AAGGGACAGGTGAGGCCACGAGCCGTGGGACCTGGAAGTGTTTAGGGCCT
GGGGGACGAGGCTGCGGCCTGCGGGCTCCGTGGTCAGGAGGCCCTCTGCC
CACTGAGCAGCTCCCACCACTGGCACACGAGCCTCTCTGGGGTCCGGCTG

FIGURE 6, CONTD.

GTCTCCGGCAGGGGTGGGCTCTGAACGTCCAGCTCCGCAGACAAATCAGA
TTCCCCCGAGCCCTGAGAAAGCCCCCTCCCCCAGCCCGTCTCCCCACCTG
TCGGTGGACAGAGTGACCCCTGCTGACCCCCTGCCCGGGCTCCCGCAGGA
GATGTGAGAGAGTAAGAGGCGGTACAGGACGGCCGGGGCGGCCCGGGCGA
GGTGCAGGTGTGTGGGTGTGAGGCTGGGCACAGGCTGGCACAGCCTCCCT
GGCCCAGTCCCTTGGGCACCTCTGGGCACCTCGGTGTGCCTGCCTCCTGA
AGGGATCCACCCTCCAGCCACCTCCTCTCGGGCCAGCCCCCACCCCACCC
CCGAGCTACAGATGCCTGCGCATTCGCCCCAAGTGTCCTGGACCCTGGAG
CCAGGCAGCCCACCCGCTCAGCCTGGCCAGACCCAGCGTTGCCCTTCACG
CCCTCCTCCCTCCCGCCGGGTCCTCGCGCTCGTCTCCTCAGGTTGGAAGC
CCCTTCCCACCTGCCATCTTGCCTGCGCCCAGGATACACCCCTCAACTCA
AGGCCTCACTCCTCGCCCTCTCCAAGGCTCTGTCCAGGCCCCTCTCTGAC
CTGGCACCACCTGCCGCCTCCTGGCAGCCCCAGCAAACCCCCTGCCACAG
TCCACGACAGTCCTCTTCTGGCTCTGCCCCCAGGATGCTTCTAGAACTGG
GGGGGGGGTCCTTCCAGCCCACGCAGCATCCACTGGGCCCTGGGCTCCCT
CCCCAGGTGCCCCTCAGAGCTTGCAGCTGGTGCAGACGGCTCTGCTCCGA
ACCCATGCTCCCTGCGCCCTTGGACCTGGTGAGATGTTGCAGGTCATTTG
GCTGCACCCAAAAGAGTGGCCCCTCAGGGTCCCCCCTGCGCCCCTCCATC

Contig 90 (350 bp)
GTACTGTAGGGCCTCATTCGAATAGCCTACTAGGTCACAGCTGATCCACA
CCTTAGGCCATCACAACTTCCCAGAGGTAGTGCCGCTCCTGTCGTTGAAC
AAGACGGTAGTGACTGCTGTGAGAGCTCAGATCTGGTGGGTCACTGACCG
AGTGTGGAACCCTGGGGAAGGCTGTGGGGTGTCCCCGGCTGGGTGGCCA
TGTCATGTGCCCCTTTCTATCCCTTGGACGAGGCTGGTTCACTCGGCTCT
AGAGCCCCAAGCCCCAGCTGCTCTGCCAACCCCCCAAGCCTGAGCCTCAT
CAGACCCACCACCCCATCGCCATGGCTACGCAGGACACACCGCTCTCCAC
CCCCACCAGCCGCCCCACCTCCCCGAGGTTCCAAAGCTTGA
Contig 91 (1464 bp)
TCCAGGACCTGATGCAGCAGCCACGTCGCGAGGCCCCTCCCACGAGGCCC
CTTGTTGACCAGCGCTAGGGAAGGGGACCAGGGAGATGCTGAGAACGGGG
CCTTCCGAGGGGGCAGGTGGGACTGACTGTGACCCAACACTCCCCACCCC
CCTCTCCCGCTCCAGAGGGTGCCAGCCTGGAAGCTGGCAAAGTCCAATCC
ACAGGTGGGCTCACGTGGGGAGGCTGGTGGCCCCCACCTGGTGGGGCCCC
AAGCTGCCTCTGGGCGGGGTGGGGGCTGCTCCCAGCAGGGTCCCATCCAG
CTTCTCCCTGGGGAGACTCACAGTTCTGGGAGAAGGGTCCTGACTGCACC
GCAGCGCCCGCCCCCTCCCCAGACTCACCCAAGTTCTCTCTCTGCATCGG
TGACTGGTCTCCGCATTTGCCCAGGCTGGGCATCTGCCCAGAGGATACGT
CCAAAGGCAGGGCAAAGCCGGGCCCGTCCCCCGGAGCTCCCCACAGGCGC
TGAGGGCTGGGCTGGATCTCGGGGGGGTGGAGGGGAGGACTCAGAAGGTG
CAGCGGGGTGGAGCGAGGCTGAGCCAAGGTGCACGCGAGGGCCAGAGAAG
GCCGAGGCGGGCAGGAGGAGAGAGCGCCAGCCTGGAGGGGGGTGGGTGCC
CTGGGCAGGTCTGGGGCTCAAGAAGAAGAGAGTGTGTGTGCAGGGGGCTG
TCCAAGCTGCCCGGGAGGCTGCCTGCCCACCTCCAGGGAGCAAAGCAGGG
AGGCTGCAGCTGGCCCGGCCGGCCGCTCTCCAGGACCACGCGTGGCCCAG
GCCTCAACGCTCCTCCCACAGCCCAGGAGACCCAGGGCACCGGGTCCATT
TACCGCGGGCTCCGGGTCCGTTTGCCTGCGCCCTGGGATGGACTGTGGGG
GCGGGGCGCTGTCTGGGGAGGAGGGAGGTGTCTGAGGCTGGACACCTTGA
AGGCAGGTGAGAGTGACAGGTCCGTGCGCAGGAGCCTTCGGCTCTGGATT
CTGGCCCTGAGCGAGGGGCTGGCTGGAAACTGGGCCGGGGCTGCCGCAGG
AGAGTGTGCAGGGAGAGGAGACGGGGTTTGGCCCCGGAGGTGCCGGGGTG
GTGCCCTGGAGTGCGGCTGAGCGGGAAGTGGGTGTTGGCGTCTGGAGACG
GGGGGTCGTGGGCTTGGGATGGTGACAAGACCCCCCAGGTGGAGGCGGCC
GCAGAGGAGGCAGAGAAGCCAGGCCCCAGCCCCACGGCGGGAGGCCTGGG
AGTCAGGAGGGACCAGCAGAGCCCTGGGCTCAGTGTCACCGGTCCTGGCA
CCTCGCCGACGGATGTCCTGGCCGTGCAGTGGTTGTCCCCTCACCCTGAG
CCCTGAGAACCATGCAGGATGCTGGTGTCACAGCAGGAGAGGGCCAGGGC
CTGGGGAGGAGTCTTACTGGAAGGCCTTCTCCTTCCGTTTGCAGCAGGCG
GGAATGACTGGGGG
Contig 92 (694 bp)
TGGAGCCAGGGCACGGCAGAGCGGTCCCGAGGCCGTGCGTGCTGACCCGG
GGGATGGGCGGACCTGGGGGTGGGCTGTGAGCCCAGGCATAGGGACCCCG

FIGURE 6, CONTD.

ACTTGGGCACGGCCAGGTGGGGCCGGGCAAGGGGGAACAAGGACGCTGGC
CTCCAAGGGCCCCACGTGGGCACAGAGGAAGAGCCGACCCAGGTTGTGGG
CGCATGGAACCCCCCACTCTGGGGGCCAGGAGGCCGAACGTCCCAAGGGC
TGAGGCTGGGAGGGAAGAGTCCCTTTGGGGGTCAGTCAGTGTCCCTTGTG
GGTGCCCCCCTGCCACTGGCGGCACCTCTGACCCCAACTCCTTGCGGGTG
GACGGTGGATGGATTTCCTGCAGCCTTTCTTCTGGAATAGTCTCTGCCAT
CCTCGGGGAAGCAGTGATTGCTCTGCCCAAGTCCAGGCCCCGCCCTGCAA
GGTGCCTCCCACCCCAATGAGCCCCCGGACAGTTCGAGGCTTCTCACGC
TACTGAGGGGTATGAACAGCTGTCCCCCTCGGAAAGTGGGGGACAGGCCC
CTGCCACTCCATCCTCGGGACGCCCGGTCTAGTCAGCACTTGTCTCCCTG
CCTTGTGCCCCCCTGACCTTTTTTGAGGACCATCAAAACCTCAGCCTCTG
CCCCAGGAGGTCAAGCCCCCCGTCCCCCAGCCCCCAGACCAGCA
Contig 93 (900 bp)
CCAGCCCCATCCCCCGGCTGGTCCCCCACCACACAGAGCCCCCGTTTCCC
AGGGGACAGCACAGCCTGCCCCCAGGTCTTACATAAAGTCACCTTCTCAG
AGCTCCTGTCGCGGCTCAGGGGAATGAATCTGACCAGCATCCATGAGGAC
ACAGGTTTGATCCCAGGCCCCGCTCAGCAGGTTAAGGATCTGGCGTTGCC
GTGAGCTGTGGTGGAGGTCGCAAGACGTGGCTCAGATCTGGTGTGGCTGT
GACTGAGGTGGCGGCCAGCAGCTGCAGCTCTGATTGGACCCCTAGCCTGG
GAACCTCCATATGCCGCGGGTGCAGCCCTGAAAGGACAAAAATAAATAAA
TAAATAAAAGAAGTAAACACACCTTCTCTAGCCATAACCACCTGCCTAGG
GGCGGAGGGCCAGGAAGCGGCACCCCCCGCCCCAGGCTGCCCGTGCGCCC
CGGGCAGGCGGCTCAGCCTGCTTTTTCTCTGTGATGTGAGCCGCCCCAGC
CCCACATGGAGGGGCTGGGCTGCGCAGTAACTGCTTTAACTGACGGGAGC
TTCGACCAGCAATTCACCAGCGGGCATGCAGCCGGGAAGGGAAGTTATTC
GTGTGTAGCTATTAGGCGCCGGAGTGAGGGTGTGCCTCGCCCTGGGCCCA
CCCCTGGGGGGAGGCATCACAGGGGTTTTGAACACCTGCCCATGAACACG
GGGCAAAAGCCAGCCAAGGGGGCAGGTGCCTGAGGCTGGGAACCAACCCG
TGTCTCTGAAATCCGGGGAATGCCCACTGCAGGCATGTTCAAAGGGTCAA
GACCGGGGCTCTGCCTGAGAAGGACTGGCGAAGGCCAACTACAAAAGCGC
ACCCCTCTGTGCAAACCCCCAACCAATGGAACAAAACTCCAGAGGGGCCA
Contig 94 (550 bp)
AGTCTGGGCTGTGTCCATGGGGTTGCCAAGGTGCCAGGCAGAGACCTTGG
GGACAAAGGTCCTGTGAGCAGAAGGACATGGCCACGTCCCCTGCTCAGCA
GGTGCCCAGGCTGGGGTCTGATGCCCTCGCTGGGGTGGGGGCGGGTTGAG
GGGCCAGGCCCAGACACCCTTCGTCCCTGCCGGAGTTGTTTGCCCTTCTG
TTCCTGGAAGGCCCCCCTGCAGGTACAGGAGGCCCCTGGGGCTGACGCTG
CACCTTCTGACACCTGTGGTCTTGGGGATGGGACAGGACAGGGACACCCC
GGGGCTGGACGGAGCGGGTAAGACAGAGAGTTGACTCTGTCCTCGAGTCT
GTGCAGGGCTGTCCCCGGCTTGGGCTTCGTCTGCAGGGCCTTTCGGGTCA
GGGTGGCCTCAAGGTGACGAAGACCTGGTCCTCGGGAGTCTGCAGGCGCA
AAAGTTGGAGCCCACCCCCCCGGGGAGGGGGCCCCAAGGACAGGAGGGCC
CAGGGAAGTCTGGGGCCTGCAAGGCCGTCCGGGCTGGGGAAGGCCAAGGT
Contig 95 (1200 bp)
GTTTGCTCTCAGCAGGCAAGGGCCTCCGAGGCCTTAATAGCCCATAATGA
CAGCGCCCGCTCCTGGCATGGGGCCCCGCCTGGCATGGGGCAGGGCAGGG
CAGAGCAAGCAGCATGCAGCTTCTACCTTCTTCCTGACCTCGTGGCCCCT
TCCGAGGCCTCAGGGGGTCCCCCGAGTGGGACCCCAGCCCTGGCTCTCCT
CTCCAGAGCCAGGCCCAAGGCTGGGAGTGGCCCAGAGATGAGGGTGCCCG
AGCAGGGCACTGCCTTGGCGTCCCCATCCCTGGCGCCTCAGGGCCGTACT
GTCCAAAACCAAAAGAAAGCAGTCAGCAAAACTTCTCCCAGCAAGCTGGG
GTCAAAGGTCGCTTCCGAGGCGTGATCAGGGTGGCCTTTGCTACTGTCAC
CGTGTGCCCTGGGAGAGGCACAGGGACACAGACACACACCTCCGAGAACC
TGGGGCTTCCAGGGCGTCAGGCTGCCTGGGCCATCCCGGGCCCCTGTGGT
CCCAGGATCTGCCGGGACCGTGAGGCCTGCGTCCCACCCTCTGCCTGGGA
CAGGCCCCACAGAGCTCACAGCCAGGGGACCGGGGACAGGGCCCCGCCTG
GGCCACCTGCCTCCAGCCTCACCCAGCCTGGGCCCCAGGCCTGTGCCTGC
GACACCCTGAGTCTCAGGACGGGCGCGGGACAAAGCCGCCCGGCCCCTCC
CCCGGCTGGGAGGAGACCCGCGTGGCCCTGACGTGTGGGCCTGTCAGAGC
TGAAATGTCACAGCAATTAGCCCTAACGAGGCCGAGGGAGGGAGCGGCGG
GGAGGCCGGCGGAGGGGATCCACGAGCCGAGGGCCCGGAGCTGGCCACCC
CACCGGTCGATTCCAGGCACTCAGGGATAATTGGGTGTTTAGAACTCAGG
CGGCAGCAGAGAGCGGGCCAGGCGGGCTGTGCCCCCCCTCCCACCGCCCC
TTAACAGGTGCCCGAACACGCAGGTCTGGGGAGATGCTGAGGTCGCCAAG

FIGURE 6, CONTD.

GGCACCCCTGGCCGTGCCGCGGTGCTATGCTGGTTCGGCACCATGGGAG
CTGCACCTGCAGCTGTATTGGTCTGTGTGTGTGTGTGTGTGCACGCGTGT
GCGTGTGTACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTACGTGG
GGGGGGGGGCAAGCCCGTGCGTGTGGTGCACAGTAGACATTTAGAAGGT
Contig 96 (600 bp)
GGGGACCAGGGCCCAGCCCTCCAGCTCCCACGCATACCTGCTAGGAGCTT
GCAACCTGCGAGAGCTTTGTGGACCCCCTGCCGGGTGACCCCTGAAGCTG
GCAGCTCTCCTTGGCTCTGCAGCGGCTCTCTACACTACCCCCTCTCCAGC
GGCCTCGGGCCCAGACATCACCCACCCGCAAGGGAAGCAGCAAGCATCCA
CCAGCTGGGCCCTTTTCCCCCAGCCTGTGACCGGCCCCGCGCCCCCTCAC
ACCTCTGCGGTCCAAGACCCCTCTCTGGCTGGGCCCTGGTGCTGCCCTTG
CCGTGCACATCTGGGGTCCATACCCCACCAACAGGCCCCACTTTTCTGTC
TCCCAGTGTCCCCCTCAGCTGCCCTGATGGGCCCACACCTGGCTTCTCTG
CTGCCCCCCTTGACCGCAAAAAGACTGGGGTCCAGGACCCCCTGCCCCAT
GACTGCCCTGGAAGACCTCAAGCCTCTCCTCTCAATCCTGACCCTTTAAG
GCTCTTGCCACGGAGAAAGCGCCTGGGGTTGGGGGAGGGTGTGGGTCCCA
AAGCAGCTTGCATACTTCTCCTGACTGGGAGCTCATTCCTCCACAGCGTG
Contig 97 (1350 bp)
CCCGCCTTATTTTTAAATTTCCGAAAACAAAAACCACACCTCTCCCGTCC
CCGAAATTATTTTGGTATAGTCTTATTCAAAGAAGTCCTGCCACTGAAGC
CCACTTGTCCTGTCCCGGGCTGCTTTGGCCAAGGGCCCTGACGGGCCCAG
GGTGGCTCATTCCCGCATCCCCGCAGAGGCCGCCTTCACATCCCATGCGG
GAGCCTGGCTTCCGGCACCCGGCTGTGCCCTCGCTGTGGCCATGGACTGC
TTTCGCAGAAGCATAGGGGCCACAACATGGGACAGCCTCGCTCTGCTCGC
TGTGGTTCCGCTGAACCTCTCAGCTGGACATCTGCGCAGCAAGCACCCCA
GCTTTGCTTCAGGCTCTGGTTCCAGGCTGGGCCCTCCTCGGCCCTGCCCG
CTGGGTGCCAAGCAGGGCTGGTCCGGCTGTGCCCCCGGGTCTATAGAAGC
CTCTGCAGGGCTTCCTACAGCCAGGCTGGGATTCGGCGGCTGCCCGGGAC
TGAGGCCCCCTCTGAGTCTGACCCCCCCATCCTTCCCTCCCACACAGCCC
CCCGCCCCCGCTTCTGCTTCAGTGAGGCCCCACCCTGCCTCACTCGCTGA
CATTTCCAGAACAGGGGGTTCCAGGAAGCCCTGAGCCTGCAGGGGACTCA
GTGACCAGCCGCATCTGAATTTTCCCTCCTTCTGATCTCTGGAGACACGT
CTGGCTCAGCCTGGCTCGAGTGCCCTGAGCTGGGGACCAGGACAGACCTG
CAGATGGAGGTCTGAGCCTGGGCAGGGCAGGGCCCAAGGCTCAGGGAGAA
ATTGCAGGTGTGAGATCAATGACCGGAGCCTGGATGGGGCCGCCCTGGCC
AGGGCAGCTTTCTCCCTGCAGCTCCCTGCCACTGTCCCCCCCAACTCTGG
GCTCCTGCTCTGGACCCAGTTCTGTGTTCCCCTCCTCCCAGCCGAGCCAC
CCTCCCCCATTCTGCCCCCCCCCAATCCAACACCCTATCGTGGGAACCAGT
GGAGCTGAAAGAAGGACCCCCCAAGGGCCCCCCAGCCGCTGTAATCCTTG
GGGGCCTCTGCCCAGGTGCCAGGTCTCGGGCAGGAGGGGCCGCGGGCACA
GCCGTGGCAGATGCGCCCCCCAAGCCTGGGCTCGGAGGAGCCCCGCCCCC
ACTGACATTTCCAGGCCGCCCGCTGCAGACCCGGCTGGCCGTGATATTTA
GACAGGGCTTATTTGCCGTGACTGGTTTTTGATGACTTTGGGGCCCAGGA
TGAGCTCAGCCGAGCCCGCGTTGGCCCACCTTGGTCTCAGCTTGGGTTTG
ATAATATAACGCGTTCAACTGAACCGCTGACGCCTGCGTGGGCCGAGGCC
Contig 98 (1354 bp)
GCTTGCAGTAGTTCATCAGATTGGACGACTCATAAATGTCAAGACATCTA
AAGATTGGTGCATCCAATCATTTCCCACCAGGTTGTTTTTTTGTAGATGT
CAAGAAGCTGACCCAAAAACTCACGTGGAAATGCACGTCAACTGGGAGAG
TTGAAACAATTTCTAAAAAGAAGAAGGACGTCGTGGGAGGACTCTTCGCG
CTCTTTGGTTTCGCTTCACTTTATATTATTAGTTACTGATTTTCCTAAAA
GCTGCAGTAGTCCAGACAGTGGGCCTCTATGAAGGGAGGGGCTCAGAGAT
GGTTGGGACAGAATAGAAAGCCCAGAAACGGACCCCCGCAAATGTGGTCA
ATTGAGTTTGGGCAAGGATGTGAAAGCGGTTCAGTGGAGAAGAGTCTTTT
CAAGAAATCTCTGGTCCTGGATCCACTGCTCATCCAGGCCCAAGAGTGAA
CTTGGCGCACATTTCTCACAGTGTATACAAAAACTGACTCAAAATAATTC
ACATACCGTCGTGTAGCGTATGAAGCCATGAAACATCCAGAAGAAAATCT
CGGTAACCTCAGGGCATCTGGGGCCTCCACCCTCAGCACCACTGGCCTTG
GGGCCAGATACTTACGTGTTCTCCTGTGCACTGTGGGACGTGCAGCCAAA
CCCCAACAAGGTGACCATCAGAAATGTCTCCAGACGTCGCCAAATAACTG
CCAGAGAGCACAGGAGCCCCTCACTGAGAACCACAGGGTGGGGCAGAGAG
ATCTCAGACATGACACGATTAGGGGAAAACAATCTGACACACTGGCTTTG
TTAAATTTAAAACTTTTCCCCTGTAAAAGGCAATGGTAAGACATTAAGAG
GCGAAGTGGCAGACTGGGAGAAAATATTTGCAAATCATGTATCAGATACC

FIGURE 6, CONTD.

AAGAAGATGCAGGAAATCCTCAAAGTTCAGTCACAAGAAAACCCAATTCA
AAAACCAGCAGAGCAGACATACGATGGCAAATAACCACGAGAAAGTCAGC
ACCCGCTGTCCCTGGGGGGACGCGAGTCAAAGCCAGGAGGACACCAGGAT
ATGCCCACTGCCAAGGCTACGGATAACGGGAAGCAAGAGACACAGACAGA
AAGGATGCTTCGGTGCTGGGGAGGGTGGGGTGGGGCGGGGGGTCCCCCCC
TGGAGCAGGATGTGAAGGCACTTGGGGGGGGCTCTGCACTCCTGGGGGCC
TTTGGCACAGGCGGAGGGCCCGGGAAGGCTCTAGGGGCACGGAGAGGGGT
GCCAGGCTTCCTTACCCAGCCCAGGCAGACCAGGCCCTGTCATGAAGCCT
GACGTGCAGCAGCAAGAGCAACATGCTACAGACATGTGTCTGTCTGTGTG
TGTG

Contig 99 (1000 bp)
GGTTCTCAGGCGCACGGGGCAGAGGCTGAGGGTCCGAGGGGCTTTGGGTG
CTGGAAAGCCTGAGTTTGAATCCCAGCTCGGTTTCTTAAAGCTGTGTCTC
CACGGCCAAGGAATGGGGCCTCTCTGGGAAAGGTCTGGGGTGAGGCTGGC
GGGACCTGCCAGCCCCGGAGGGCATCTGACCAGACAGCTTCTCAAGCTCA
CAGGGCTTCATGGCAGGATGGGGAAGGCTGTGGTGGGGAGTGGGGAGCAC
TCGACACCCTGTCCAGGCCTCTTGAGTCACGGTGGCCTCTGAAAAGGGGT
TCTCTGTGTCCAATGAGCAAGTCTTTGTCCGGGGCAGGATTACTAAGTCC
AAGGGTGTCTGCCCCTCCGTGGGGCACAGAGCAGGGGCCCCAGATCACGT
GGCTGTAACTGCCAGGTTGCAAAGCCTGCCACCATGTCCCACTGGGTTCT
CCAGTTACCTTGGGAGGTGCAGGGTGGGGTGATGGGGAAACTGAGGCAGA
GAGCTGGCAAAAGAGTGCCGGCAGGGACTGCGGGCGCCAGACCCAGCTAA
CCGACCCTCACACGGAGCTGCTTCTACTTTGCAGCCTGGACGTGGGAAAA
GGTTACCCCACAGCAGCGTGTGCAGGCACGCTGGTATGTCTGTGTACTTA
TGCATATGTTCTACGTGCATGCACGTGAGTGTGCTGTGTGCATTGTGCCT
GTGTGTGTGTGCATGTGTGTGTGCACTCATGTGTCTATACGTGTGTGTAG
TGAATGCTTGTGCATGTGTATTTGCATGTGTATGTTTGTACGTGTGCAGT
GAATGCATGTGTGTGCAGTGGCGGCATGTGCGTGTGTGCGCATGTGTCTG
TTTATACCTGTGTGTAGTGAATGCATGTGCATGTGTGTGTTTACATGTGC
ACGTGAGAATGTGCACTCGTGCATGTTTGCATGTGAGTTTCATGTACACA
TGCTTTTAACGTGTGCACGTGTGCACATGTGTTTCTGTGTCCCTTGCACG

Contig 100 (1500 bp)
CGTATAAATATATTAATATAGAATAAAATAGATTGATAATATAGATAAAC
TAAACCCATTATCAATACCGGGTGGCCCCAGCAAAGGATACTAGCCAGTT
TATCAAGGTGCTAAGTCAGCACATAGAATGGCCACAAACGAAAACCTGTA
CTGCCTATGTCCACTCTAATGGAGTATGCCACTGACATCAGTGGTAGGTG
AGCTGAGTCCATCTGGGCTCCCAGTTCGGCCCCGGCTTGTCCCCCAACGG
AGGTTCCTTCCAGGGTTCCCCAAACCCAACCGGGCCCCCAGGTCTCCCTG
TCTTGACTCGTTTCTGGAGTCTTCTGGGGCTCTGCAGTCCTCCCTTGTTG
GGGCTTCTGTCCCCCTGCCCCTGGCCTTGCGGGCTCGGCCCTGCCCTGGG
TCCCGGGCCTGCGGGCTCACCCTCCTTCTTTCCCTGGAAGAGAGGGAGCC
AGGCTGGGCCGGGCCAGGAGGGAATGCGCCTGACTCTGCTCCAGATGGAC
AGGTCGGGACATGCAGTGGCCTCGCCTTGGGCTGCTGAGCCAAGAGCAGG
ACGGGTTCTTTCTGGAATGTGGGGCCAGCCAGGTTCAGCGTGTGGGTGGG
CAGCCGCCAGCATCTGTCAGGGCCGCTGCAGGCGCGGGGAATGACCTCGA
CTTCTGCTTGGCACCCAGCTCTGGAACAGCCCCCTGCGGAGCCTCCGCCC
AGAGCTGGGCCAGAGGGTCCCCTGTGCCGGGGACCCCAGCAGGGCCCCTC
CCTGACTCTCCAACCCACCTGCCTGGGAGGAGTGGCCCCCTGGCCTCCGT
GGATCTCTGGGTCGGGGCTCAGCCGGCTTGACAGCCTGGGAACAGCCAAT
GCACATCCCCAGGCCTGGCCACACCCTTCCACCGGGAGCGGGCGGATCTG
CATTTCGCCAGGCTCTGCGGGCAGCTCTGAGAGCCCCGGGTCTCGGAGCC
CAGCCGTGGCCGTTGTACGCCCTGGGGGCTGTGGACAGCGTGTCCTCATT
GCCCCTCCGAGGTCCGGCCCAGGTCCCCTCCCACCTGCTCGCCCAGAGCC
CTCTCCCCACCAACCACACTTCCTGCTGTTCTGCAAGCGGGACACACACT
CCGGTTTCAGGACCTTTGCACGTGCCGCTTCCTCTGCAGAGAAATGCCTG
GAGCAGATGTTTGTCCGCACGGCTGCTCCGCGAGGCCTACCGAGAGCCCC
TCACCTAAACGGCCGGGCCTCAGCAGCCCGGGGCCCTGTCCCCACCGCCC
AGGTGGTGGGTTCTCCTGTGCCAGTGTGGGCATCTCTGTAAGATACCTGT
TTATCTGCTCATCGTCTGGTCTCCCCCAGAAGGTAGAGCAGGGCCCCGGCA
CAGCCGTCCTCGGGGTGGCCACTCGCCCTTGGGGCTCAGCCTCCATGCAG
GGAGGGACGCCTGGTGACACGAGAGCCCCGTGTGAGTGTGCCGGGCCGCC
AGCCTGCCTTAGGTCACAGCCAAAGCCGGCATTAACCACCAGGCCCTCGA

FIGURE 6, CONTD.

Contig 101 (600 bp)
TCTAGAATACCTGGCCCTCCAGGGACGTGTCCTGTAGCTGCGGCTTTCAG
GGCAAAGTGTAATTAAACATCCCCAGGCTTCCCTTCCAGTTGGCACAGGG
CACCCACATGAGGAGCAGCCTCTGGGTGCCAAAGGGCCCACTGGTGCCAG
GCGCTGGGCTGAGTGCACCCCCGCATGCTTCCCGCCCACTCACCTGCTGG
CCCCACCCCTGACCACAGCACCTGTGGGAACACTAGGCCTGGCAGCCACA
CGCTGCTCTCACTGGAGGCCAGTGCCAGGCAGCCTGCTTGGCTACGCTAG
CAGATGCCCGCTCGCCTCTGCCCCTGCCCCTAGCCCATGCAGGAGCCCAG
GGTGGGGCACAGGAAGGACGATTGGGGCCCCAGGTCAGGCACATCCAGGC
CACAGCCGTGGCCACACGAAGGCGGCCCTGAGGGGGCGTTGGGGGGCAGA
CCCTGCCCCCCCGCTGCCGCCCCAGCTCCAGGCATTAATTCCCAGGGACC
TGTTGCACTGGGTGGCCGCCAGCCTGCCCCCTTGCCTTCCAAGGCCTCTA
AAATGCCCCTCTTTTCGTAAACTAGGACTTACCAAGCTCAGCGAGCCCTC

Contig 102 (1867 bp)
AGTATATCGGGTGAGACTGGGGACCGGTCTGCCGGGAAGCCCCACCATAA
AGGCCACGTTGGGCCACAGTCCGGGCCACGTGAGTGTGGGCGGGTCCGCG
GGTCTGCTCTTGGAACACCAGGATCTCTAAGAGGTACCAGCCGAGGCCAA
GTTCACGTGAGCAAGTGAGCAAATGACTGAATGAGAGCGTGAGCGAATGA
GTGAGGGGTGAGTCCGTCCACCACGCAGCCTAGGCTCAGCCAACCGCTGT
CCCCGCGTCTCCACTGGTGACCAGAACGGAAAGAGTGGGGAAAGACTGGT
TGTCTCCCACAACCCAGTCCCCAACCCCCCTGGACGCCCCACCCCTCCAG
GGGTGCCGGGCCTGGCCTGTGGGCCCCAGTCTGGAGGCTCTGGCACCTTC
CTCATCCGTTCTCCCAGCACCCCAGGTTCGTGCTGAGCCCTCCTGGCCCA
CAGGCCTCGGGGACAAAGAGGGCCACCTGGAGGCTCAGGGAGCCTCACCT
GCCTCGTGGTCCTGGCGGAGGCGGGTCTGGACATGTGATAGACCGGCCTG
GGCTCAGCAGCTCCTGCTGGAAGATGTCAGGGACAGCCTGGGCCACTCTC
CCACCAGGAGAACTTATTCCTCGGTGGGGTCCCCCCGGGGAAGGGATGGG
ATCCCAGCGGGGACCCCAGAGCGTCCAGCACACGGACCTGTCCCTCCAGC
CCCTGCCCCACACGGATGCTCACAGCTCAGCCTCGAACACGCACCTGTTG
GACTTTGCCTCCTGAGGCTGTCTTCTCAGCCGACGCGGCCTCCGCTGCA
TGGTCTGGAAGCCCAGTGGGACTCGGTGGTGACAGGGAACAGGGGCTCTT
GGAGTGGGGTGCCGGGGGAGCCCCGAGGGAGCTGCTTGGGCCTTTGATGG
CTGAGTGGGCTGAAGTCAGGCAGGCTCCCCCAGGGCTCCCTGACCCCCCC
CACCTCAAAAAATCCAGAGCATCCTTTGCTTTGGGTCTGGTGAGGCTCTC
TGAGGTCAGACCCTGCGTGGCTCGGCCAGTGGGGCTGGAGCAGGAAGAAA
GCAGGACAGCCCCCGCCCCTGGCCCAGACTCCCCAAACCCAGCAGGAGAC
ACCTGAAACGGGATGGAACCATCCTGAAAAGAGCCACCTCCTCCTCCTTA
TGCATCAGCTGCCGGGGTCTGGGGGCCCGCCCCAGGCCCCAGATGTCCGG
GCTGCTCCCGTCTCACATCCAGGGGTTTCTGGGCCCAGGACTCTGTCCCC
CCAAGCATGCAGAGGGTCCAGGCTGGGGTCTTCATGCCTGCCCGTGTGCA
TGGTGGGGAAGGAAGGGGACAGTCTGGAGACCCCCCGCCCTCCCCATGCG
TGGCGCCGGGGGACAAAGCCGGCTGGGGTCTCAGGTTTGGGTTCAGAGCA
AACGTTGATCTGACCTGGTTCTGAGATGCTCGGCCCGATGCTGCGTTGTC
CGCTCGCATTTTCCTGTTTTCTCTGGGAGGCCCTGCGTGCGCTGTGGCTT
CCGGCCAGCCCCACGGAGGGACGCAGGGTGGCTGGCGGGGTCTGGGGGCC
CCTGCCCGCACCAGAACGTCTGGCTCAGGTTTTTGTCCTCGTGACCCATC
ACTAAGGGCCACCCTCTGACCCGGAGCCCTGTCTCCGAGGTGGGAATTGG
GGGCTGTCCCTGGCGTCATAGGACCTGGTTGGGGGCATCCAGGGCTGTGT
CATGCCCCTCCCCAGAAGACTCTGGGGGCTGCGGGAGGGTTTCCCCAGCT
TCGGGCCAGCCTGGGGAGGGCGGAAGGCGCTGGAGGCCTTGCCTGTCCCA
GGGAGCATGGCTTCGCTGCAGACTGGGGCCCCGCACACCCAGCCACCACT
GGCCGTCTGGAAGCACT

Contig 103 (650 bp)
GTTGAGGATTCCTCGGCAATTTCCTCGTCACTGGCGCTCCAATCGCCTCG
ATGGGCTTCTCCTCCAGATACAGCTGCAGATCCTGGGCGGGCACACCGTT
GAGCGTCACCTCGTAGTGCAGATTGCACTCGTTGTCAATGGACATCCAGG
CCATGCCGACGGCATGTGGATTCTGTGCATCCGTGTGCTCCTGTCGCTTC
AGCAGAATGGGTTCCGCCGAGTCCCGAGCATCGGCCACTGGACGGGGCAC
TAGGCGGCCACGGATCAGGCTCGTCTCATGCTCGGTGGCCACATTAACGC
CCAGTTCGCCGGCATACAGCGACTCGAGGACCTTGGGACCCAACTTCTCC
ACACTACCAATGGCCTGGTTGAAGTTGAAGCTCGGCGTCAGATCCTCCAG
CTTGGCCTTCCGCTTGCCCTGCTCCTCAATCAAACTGATGTTGGGCCTAT
CCCGGGTGTTCACGTGCTCCGTTTCGATGTTGTAGGCCAGAGATCCATCG
GTGTTCAAGTAGACCCACGCCAAACCGCTGCTCTTGGTCGAGGATTCGGC

FIGURE 6, CONTD.

ACTGTGCGGCGCCAGCAGGGTCTGGAAGATTTCGCAGCTGGCTCGGGTCA
CGATGTGTCCCTGGATGCGCAGATGTGGGTACTTCTTGGACTCCACGGTC
Contig 104 (1630 bp)
GGTGTTGTCACTGCTGTGGCTCAGACCCCTGCTGTGGCACAGGGTCCATC
CTTAGCCCAGAAACTTGCACATGCCACAGGTGCAGCCAAAAGAAAATTCT
TACTAATAAGTTGTTCATTTGCCTTTACGTAGAGTGGCATCAAACAGCAA
ATTTAAAACACCATCTATCAATACATAGACCGCGGTCAAAGGGAAAGAAC
TTTCTATTTCAGCACCTTTAACATGGCTTTGCCCGAATTTGGGACCAGGG
TGCTGTGTTTTCATCTCTCCCTGCAGGTGGTCCCCAGATGACCAGGCCGG
TCCTGGGCGGGACGAGCCGGACTGTGGATCCAGTTGCTTCCCAAGACAGG
CTGACAGGAGAGCAGCAAGGGCCACCCCCAACCGAAACCAAAGCCAGAAC
GAGCAGAAAGATGCCGTCTTCCAAGTGGGGGCTGGGAGCTTCCTCCCATC
CTCCGGAGCCGTGAGGCTGCCCTGGAGCTGGCAGGAGCCACAGAGGACCC
GGCTTTGACCGCCCCTCTGGGACCCACAATCAGGACCCTGACTCAGATGC
TGAGGGGCCTGGACAACACCCCAGGACCCTGCTGCTTCCCCAGAACCGCT
GTGTCCATCAAGGTCCAGATGGCACCCGTGTCCCCACTGGAGCACGCACT
CCGTGGGGCAGGCTTTCCCTTGGGCACCGATGCACCTTGAGGGCAGAGAC
GGGGCCCAATAAACGTTTCCAAACCAGTGGGTGAGGGACCCGACCGGCCC
GACACGGCAGCCCGGATGCAGGGACTCCGTGCTTGGCCCAGCCTCCCTTG
GGGTGCTCCTGTGTCCTCAGGGGTGGATAGGCCATCATGTGGGTGGCCTC
TGGGGACATCCGTTCTCTGATTGGGTGAGTTTCAGCCACAGAGATATTCC
CAGGACTACAAAGCTGGGTCCCTTGGGGCACCTGCTGTCACAAAAAGACA
AGGCCCTGACCCCCAGTAGCCAAGTTCCCCCAGGGGCTCCCCAGGGTCTG
GTCATCCAGACTGTGCCAGCCGTGCTGCCCGCCCCAGTCCTGCCTGACCC
GAGTCTCTGTAAACATCCCCCGGCCCCACCCAGCTTTACCCCAAGGCCGA
AAGCACCAGCCCCCCTGCACCACAGATGAGGCCCCCATGGCTCCCCGACC
TAACTTCTGTCTGCAGTTGGCTTTCAGCCTCGGGTGGGGGCAAGGCCTGC
ATCTCAGGCTCCCGGGAGAAGTTGCTGCCTCCACAGCAGAGCCAGGGGCC
TGCTGACCACCTGGGCCGGGTCGGATCTGGTCTAGAATGCTGCTAAGGTG
TCCTTGCAGGCAGCCCCGGGCGGCCCCGCCCTCCAGGAAGGAAGGGGACA
TTGCCAGGACTCAGGAATGAAGCCATCCCAGGTTTTGAATCCCCGGTCCC
ACCACCTTCCACCTCTGACCTCAGGCACCTCGGCTTTCAGAGCTGCCCTT
TCTGACTCTGGGACACGGGGCTGTGAGGCGCTCTCGCTGTGTGACAGCTG
GGGGGGGGCACTCTCTAACGAGGGTGGGCGTGCCCAGGTGACTGACCACA
GCCCTTTCCTCTCTCAAAAACGCCCGCCCGAGTGACCTCACGGAGGCAG
GGCCAGGAACCCCAAACCAAACCAGAATCA
Contig 105 (1820 bp)
AGTGAGCCCTGCAGGACAGTCTGCTGAGGGGTGTCTGGGCTCCTCAGAGG
CTCATGGCCACGGGCACTGGGAGGATAGCAGGTGGACCCCTGCATCCAGG
TCCCAGGTCCCAGGTCCCAGACCCCCGGACAGGCTTTCTATCTGCAGGAG
GGGGGCTCCTGGGGCAGCAGGGATGTGGCTGTGAGGCCTCGTCAGTCTCC
CTGTTTCTATCTCTCTCTGTATCACACACACACACACACACACACACACA
CACACACACACGCACGCACGCACACACACAGAGGCGTGACCAGGGCTGCA
GACAGGGCCATGGGAGGACTGCCCGGCAGTGCACCCAGATGGCCACACGG
TGGGGCCCTCGTCCCACTTTTGCTGCTGATGCTTCCGCCCAGGCTGCTGG
GAGCAAGCACTAGCTTCCCAGGGCTCTGACCAGAGAGGGATGGGAGGGGT
CATGGGTCAACAGGCGCCAGGGAATGGGGAATAGGATCTGAGGGGCGGGG
GCAAGGGGCCCAGGCGAGGCTGCAGTGCCCAGAGCTCCCTGCACCTGCAG
GACCAGCCACAGGCCAACAGCTGCAGGCAGAGCAGGGCTGCTCCTGTCCC
CAGAAGCTGGCACAGCACATGGGGTCTGACAGCCCCACCCCGGGCCTCCC
ACAGAGGGGCGGGTCCCCCAAACTCCTCCCCCGTCCCACCTCACAGCTCA
GCATCTCCACTGCCTGAGGACGAGCCCAACACACGGGCACACACACACAT
GCACGCACACACATGAATGCACCTGCAAGCACACACTCACACGTAAGCAG
GTACACACATGCATGCACACAATGAACACACATGCACGCACACACGCATG
CACACACGCACACACACTCAAACACGTACATGCAAGCACATGCTGGTCCT
TTGTCCCCGTGGAGGGGAGGATGGAGGCCCAGCCCGTGGGGAGGGCATGT
GGAGTGTTGGGGGGCTGGCTCCAACGCCCTCGCTCAACAGGCACCAACGC
TGGACTGAGATAAGCCGGGGCGCTGGCTCCCTTGGGGCCGCTCAGCAGGT
TTGACGCCCACCACAGGTGGCACTGCCTCTTTCAGAAGACGGATGTGGCC
ATGCCACCCTCACAGCCTCACCAGTCCCCCCTCAGCTTTAGTGCTGTCCC
TGTCACTGTACCCGGGGCCTTCCTTCTTCCAGGGCCAAAAGCGAGTTCAG
GGGACAGTGGCGCCCCCATAATTACTCACCCAGGGTGCTGTCCTCTGTGG
TGGCCTTGAGGCCAAGGTGCTCCCATGGGGGCCCACAGGGCTGGCAGGGT
CACTTCCTGAGAGCACCCAGGGCCAGGGGGGTGGCCCAGGCCTGGCCGGT

FIGURE 6, CONTD.

CCCCATCTGGAATGAGGGCCTTGCGCAGAGGCGGTGCACCCCTCTTTACA
GCAGCCCCGGGGGAGAGTGACTCCTGCGTCATGGACCTGGGGGCTGACCT
GTCACGTGTCTCGCCCAGTTGCACCCCATCCATTTCCGGGTGGAAGGGAC
AAAGCCATCCTGGTCGTCTCAGAGGACCTCTGGAGCCTCTTGGCCCCAGC
AGCCCAGCCCCTCCCGGGCCCGCATCCTCTGCCCACCCAAAATCACCTGT
GCCCACAGGGTCCCCTTCTGGGTGTCCAGGGCGACCCAGAACTGCCCCTG
CAGACACACCCAGCCCAGGACATGGCCGCCTTGCCGGGCCTGTCTGCCTG
GGGCAGCCTGACTGCCACAGACAGGCCGCTTGGAGGACCATCTGCCTGAG
CCCCCAAGGCACATCCCACGGGGCCCACACAGCCAGCGCCTGTAGACGAT
GCCACTTGGGGTGGGGGGAG

Contig 106 (1500 bp)

TGCCGAATAGAGGTGGAAACCAAGACCCGAAAAAATGTCCACATTTTTCA
ATTATTAGAAATTTAGAAAAATATTTACAGGAGTTAAAAGGTATTCCAT
TCTGGGGGCGGGTGGGCATGCCCACGGCATGCAGGCATTCCCCGACCAGC
GACTGAACTCGAGCCACGGCAGTCACCATGCTGGATCCTTAACCTGCTGA
GCCCCTGGGCAACTCCAGACACTCCATATTCATGTAAACTATTTTTTAAC
CAAAAAAATGACAAAGCTTTTCAAAACAAAACACATTTCATGGGAAGAGT
GGCATTGCTTCACGCCTGGATGGTCGCTGCGGCTTGCGGGACGACGAGGG
CCCCCGCGGGAGCGCCTCCGCACGGCGCATCAGGACGTGGTGTCCAGGGA
AGCGGGGTCACTTCACGGCCTCTCGGGTGCGCGTGGGTTTCCTTTTCGGC
ACCACACCCGGACTCAGCACTTGGGGGTTCTTAAACGTGAGAGGCACTGC
GGGGCTCGAAGCCACATCACTGACCTCCTCAGACTCTGTTATGTGAAAAC
CCATCCGTCCACGAGACCAAAGAGACAGACGAACAAACGCAAGGTGGCGC
CTAGGTTGGGCACAGCATGAGGGCAGAGCGGAAACCTTCGCGAAATCCCG
GCGAAGCCTGGACGTCGCCAGCTCTTACTTGACGCAAACATAGGGGGATT
CAGGAACTCTCTTTACCGCATTTGCAATTAATTTGCTGCAAATCTAAAAT
CGTTCCAAGCACAATGCTCACTGCATGGAAAAACCCAGGGGTAGGTCTCG
CCCGATCAGGATGTTTTCCCGTGCCCTCTGTGCGGGTGCTGCCCCCTGCG
CTGGTCAGTGAGAAGTGTCCCTCCACCGACGACATGAAACTTCCCAGGTC
CACGCTCTCTGCTGTCCTGGACGAAAACTCATCTCTGTGAATCTCCCGCC
AGCTCCGCGGGAGCCTTCCAGGGCTGGAAGGACGGCCGTCCCGTTCCAGG
GGGCAGGTGCACGCTTCCCAAAGCTCCGCGTCCTGCTAGGACGCTCAGAC
GGCATCACCCACAAACCCCCACGAACTGTTTCCCTCGAGGCGACAGGCTCG
CCCTTCTCCGAGAAAGCAGCCCGCACACGTCAGCAAGGGGCCAGCTGCGT
TTGTAACTCAAATGGCCACATAGAGTTTGTCCTGGAGGCACGGGGTCTGT
CTGGGCCGCACCACTGCACACGCAGAATATGCTGGGACACGCTCCGGGGT
CCAGCTTCATGGAATTAATAAAGTTTACTGCTTCACCAAGTACATTCTTA
AGTGTAGCTGGCCGCCAGCCTGGGCGTCCGCTCCGAGGCTGCCTCTCTGC
CTGGAACCCTTGTGCTGGGGGACCCTCTCTCCAGCCCCACCCCAGCCCCG
AGCCCAGGCAACATCCTTCTTGTAAGACACCCGCTACCCTGCCCTCCCGC
TTCTCCTTCTCTGGATCCAATCTCCTCCGCTTCTAAGCTCTCTTGAGGCT

Contig 107 (550 bp)

ATGGCACTCGCGGTTGTGACTGAGCTACCGGACGGCGCGAGCAGGGCCAC
GAGGGCGACAAGCGCGGGGCTGAGAACCTGTGCGAGGGCAGGTCCCTGCG
GCTGCAGACAAGCCTCTATCGCAGGCCCACAGACAGGAGCCCCCGTGTGA
CCCTCAGGCTGCGAGACCAAAGTCACGGCTCTGCTGGGAAAACCTCGAAC
CTGATCACTGGTGGGTGACCCCAGGACCTTGAATTCCGGCCTCTGCAGA
ACGCTCTGAGCCTACGGGAGTGGCCACCCTCTCGGTTAGGGCCTGTGTCC
TTCCCTGGCTTCCAGCCTAGAGCAAAAGCATTAAATCACAGTGTGGCCCA
GCCCGGACCGTGCAGGACCTTAGACAAAAGAGGAGGGAGAGAGAGATGAG
GCAGAGAGGCAGAGAGACAGAGGTGGAGAGACAGATAGACAGAGACAGAG
GCAGAGAGAGAGACAGACAGACAGAGACAGAGGCGGAGAGACAGACAGAG
ACAGAGGTGGAGAGACAGGCAGACAGAGACAGAGGCCGAGAGAGAGACAG

Contig 108 (900 bp)

TTTCTAAACTCTCTTACTAGTTCTAGTTTCTATTGTTTTCTGGGGGGGT
TCTATATAAACATTCGTGTCGTGATTGGAGATGGTTTTGTTTTTTCCTCT
CCAAACTGTATGCCATGTGTTTCTTTTTCTTGTCTTATCACACTGGCTAG
GACTTCCAGTAAAACACTAGATATGAACAATGAGAGGAGAGCCAGGCCTT
CTTCTCAGTCTTGGAGGAAACAGTCAGTCTTTCCTCATTTAGAATGAGAG
CTTTTCTTTTCTTTTCTTTCTTTCTTTCTTTTTTTTTTTTTAATAGGTT
AAGGAACTTCTCTTGTATTCTTATTTTTTTAGAGTTGTTATTTTTTTTTT
CTCTCTTTTTAGGGCTGCACCCGAGGCATATGGAGGTTCTAAGGCTGGGG
TCGAATTGGAGCTACAGTCGATGGCCTACGCCACAGCAATGTGAGATCTG
AGCCACATCTGCGACCTATACCACAGCTCACAGCAATGTCAGATGGTTAA

FIGURE 6, CONTD.

CCCACTGAACAAGGCCAGGGATTGAGCCCGCATCCTCATGGATGCCAGTC
AGTTTCGTGACCGCTGAGCCATGAAGGGAACTTCCAATAATGCACCAATT
TTAAATGAAAAAGACAAAGCATCCAGCCCACAGCCTGAGTAAGGAGTTTG
GAGGCCTGACCCCTGCGTGGTCCTGGGCCTGGGCCTGGGCTGGTCGGGGT
GGGGGGGGGTGGGGGGGACCCTGTGGACCCTCCCTCCTCAGCCAGGCCTG
CCCCTCCATCCCTAGCTGTCGGGGGCTCGGAGGAAGGCGGGTGGATGACG
GTCCCTGGGACCCCTCCTCATATGTATCTGGGTCCCTGGTCCCTCTGAGG
CCCAGGTCAGGTCATGGGAGTCAAAGGTCAGCCAAGGGGGTAGCCCAGAG
Contig 109 (950 bp).
TAACCCACTGACCGAGGCCAGGGATCAAACCTGCAACCTCATGCTTCCTA
GTCGGTTCGGTAACCACTGCGCCACAACGGGAACTCCTTTGCTTTTGTTT
TTAGGATTTCACATACACGTGATAACGTGCCGTATTTATCTTTCTCATCT
GAATTATTTCACTTAGCCTAAGCCCTTCAGGGTCCATCCATGGTGCTGGG
AGTGGCAGGATTTGCTTCTTTTTTTTTTTTTTTTTGTGGCTGAAAATCAG
TCCAGGATTATCTTCTTTTTCTCTTCATCTGTGGAGGACACAGGCTGCGT
CCGTGTGACGCTCTGCCGGGAATACGGGGGCCGATCGCTTTCTGAGCCAG
TGTTCTCATTTTCTTGGGAGAAGTACCCGGAGTGGAACGGCTGGGTCGTC
CTGCAGTTCTGTGCTGCATTTTTTGAAGACGCTCGGAGCGCTTTCCACAG
TGGCTGCACCGACTGACATTCCCACCGAAGTGCACGGATTTCCCCATCCT
TTTTCCACGTTTTCCCCGCACTTGCTATTTTTGCCCTGTGGATGTCGGCC
TCTCCGTCAGGTGTGAGGGGAGTCTCCGTGCGGCCCAGGCGAGGAGCGAC
CGTGAGCGTCGTTTCACGTTCCTGTTGGGCCACCTGCGTGGCTTCTCCGG
AAAAAGGGCTGTTCAGGCTTCTTGCCCATTTCTCAGTCTGATTGTTTGGG
GGGTTTGCTGTTGAGTTGTGTGAGTTCCGCACGTATGGGGGGCATCAACC
CTTTATCAGCTATGCGATTGGCAAGTCCGTTCTCCCATGTTCCGCCGGCC
GCCTTGGCACGTGTGGGCGGTCTCCTTGGCTCTTCCTTGGTGCAGAAGGC
TTCGGTCTGATGTGGGCCCATTTGTTTATCTTCTTTTCTTTCCTCACCGT
TGTTTTGATGTCAGATGCAAAAATCCATTGCCAGGGTCTGTGCCGAGAAC
Contig 110 (306 bp)
CGCCACCTCAATCGCCGGTTTGTTCTGCAACACGGTCCAGATAACCAGCG
CACCTAACAGGTCGAACACTGCCAGAACTGCGAACAGCGGGCTGAAGCCG
ATGGTGTCAGCCAGTGCACCGACAACCAGCGCAAACAGCGTACTTGCCAG
CCATGCGGACATCCCGGTTAAACCGTTTGCCGTTGCCACTTCGTTACGAC
CAAACACATCGGAAGAGAGCGTAATCAGCGCGCCAGACAGTGCCTGGTGG
GCAAAACCACCGATACACAGCAGCATAATTGCGACATACGGGTTGGTGAA
CAGGCC
Contig 111 (800 bp)
GTTTTCCATGATGCACCAGGGGGGCCGGGACCGCAGCAGGGAAGGCTCCA
TCCTGGCTCTGTAAGACCTTGAAAACACCTCATTCCTCTGGTCTTGGCCT
GCTCTTCGGTACGCCAAGTTGCTGAGACTGATGTGGGGATCAGTGGGGAG
CAGGAATCTTTCTGATTCAGCCGTTTCAAAGTGTCCCAAGCAGAAGCTGT
GATGGCAATGCCAAGGCTATCCATGGAGGTGGCTGTGCCAGGGGCCCCAT
TTCCTGGGAGCCCATTCCAGGAAAGGAATCTTGTAGCCCCAGGCTCCAGC
AGCCAGTGCACGGCCCCTGGGACTATCCGGGTAGATCAGAGGGAGGAACA
GAGCTGTGGATGGTAAGCAGGTGGCCCAAGTCCAATTTATGTCTGTGGTC
CCAGCAGGGTGCCCAGGAGGCCCCTCGTAACTCTTAAGAATCTTGGTCTG
GTCAGCTAAATTGTATGACCATTGTACTGAGCACACATCCCGTTTAAGTA
GAATTTTCAAGGATGACTAGGAGTTTGCCACCTGAAGGCAGGAAGGGCAT
TCCAGGCAGAGGGTACAGAGGTGAGAGGGAGGCTCTGACACTTTGGGCGT
GCAGGGGGTTTGATGTGACTGCAGCTGGCACACACAGTGTATGCCCAGGCCT
GGCACGGCTGTGTTGGTGTTTGGAGAGGAAGGGAGAGGTGAGTTGAGCCC
AAGGTCTTCCAGGCCAAAAGACTGAAGGTGACCGCGGCTGTCCGGGGCTG
GCCCGCAGACCAGGAGGGAGCAGGTGGGAGCTGGCTCTTGTTCCGGGGAC
Contig 112 (3062 bp)
CACACCCCAGGAGAGGAAAGACCCACACAGTCCTGATGACAGCTTGGCTC
GGGGCTGGAGCCCCGAGTTATAAATGTCCATCACGAGCTGTGTTCTGTCA
GAGCCATCAGTGGGAAGGCCAGGCCAGCTCAGCAGCCCAAAAATGAAGAG
CTAGGTCTGGGATTGGGCCCAAGCAGAGGGCACAGGAAAGCCACATAAAC
AAGGCACCCAACCCCCCTGTCATCCACCAATGTCACATTCAGGTCACACC
CCTGGTCTTCGGGGGAGGTCCCCTAAGATCCGGTGGCAGGGGGAGGAAAA
GTCTGACTGGATTCCTTGACAGGTGTATCAGCGGAAGGCCAGGAGGAGTG
CTCGGGCACTGCCACCTCCCAGGGGCATGATGGTCATGGACCAGATGGCA
GTTATGGGAGGAACCTCCCCCGTGGTCAGAGCTCTGGGTGCTGTACCTGG
TCATGCATTTCGAGTGGAAGGAAAAGAAAACATACAACTCCACCCCCAGC

FIGURE 6, CONTD.

AGCTTTAGGCTGTTGGTCTAAAGGTCCTGCCTCCTGGAAGAGACACGCCT
CTGTCAGCGGACACTGCTAAACCTAAAGGAAGAACTGCCACCTGGTCACG
GGACTTCCTAGGCCAACCAACCTACAGGTGACGGCCCGGAGCATCACGAG
GAGGTAGGGGACGGGAAGGGATGCATTTGCTGCTCAGCGGATCCACTGGG
GCGTTTCTGGAGCCCCCACGCCCACACTTTACTGCAAATGCACAAGCCCC
AGGCAGCAGGACAAGTCACAGTAGCTCTGGGTTATCCAAGGAGTCAGGGA
CCTACCTGGAAGAGTCTAGAACAGGTGACAGAGGAGGGAGAGGATGGTAC
CAGCAGTATAGGGAGAATCAGAAATCTGACCCACCCTGGGGGCCTGACTG
ACTCCCAGACCAAATGCCACACTCAGGTTCCCCGTCTGCCTGCACTTCCA
GGGCTGGGCCACGGGAGTTATGGCCCCAGGTAGCATCAGAGGCTCCCAG
GTACAGGCACAAGCAGCAACCACAGGAGGGATCCAGGCCAGGGAGCATCC
AAGAAGCAGCAGAAGCTCCACCTTAGGTACAGTTCTGGCACCTCCAAGTT
GAGAACATGTCCTAGACAGTGCCTGACCCCAACCCAATGGAGTGTCTGGG
ACTAGACTAGGCACGCCATTTTGGTCCCAGGTTGCCCCATCTGTACAAAG
GGTGTGCGGCCCCCAGGGGGACACAATGAGCTCCCATGGGAAGGGTCTTG
CGAATCTCCTTAGAAGCAGATGTAAGAGGTGACGTCCAGCTTGTGCCTGG
GATGTAGAAGTGGAAAAAGCACCCCTCCCCCGACAAGGATGAAAGCAAGA
GGCACAAAACAACCTGAAATTCCCAACGCCCCTGGAGATCCTTGGAGAAC
TGGGATTCTCCACCTGTAGGGGCACCTGTGAGGAGAGGCTGTGTGAGCAC
CTGCTGACCTGGCACAGAGGATGCCCAATACTAAGAAGCATCAGCTAAAA
GTCTCCAGGAATTCCTGGAAGCTGAGGAAGGGCTCAGGAGAGGGTACAGA
AGCCCTGGGGCTATAGATATAAGGGACGTGCACACCCACTTGCAGGTCCC
CATGGACCCCAGGGACATTCACAGTGATGGGCAAGATTCCCAAAATGCAC
CCCTTGTGTGTGGGCCTCGTTCGGTGGGTCAGCAGACACCACACCAAAGG
CACAAAGCACACACCCTCAGGCTACTCTCCTCCCTCTCCCTTGTGGAACA
TGAGCCTTGAGATGCTGGGGCACGTGAAAAACACTGTCACACTTAGGTCC
TGGTGAAAACTGACTGCGCCCAGCGGAAAGAATCATAAAGACCCTACACC
CACACACAGCCTTAATTACAGCTGTGAGTGGGGCTGGAGCCCCAAGAATG
TCTACACCCATAAGACATAGCGTTAATCAGAAAAACAAGAACAGCCCCAA
CCCCACCACCAGGCTGACAACTAACAGGTCATGTTGGAATATCACTGGGA
ATGTTCTAGGAGTGTAGAAAGACACACCAACTAGGGCATGATGCAAAGAT
AATACTTCAGCCTGGGAGTGGATGTGACACAGGGAAAAGCATAAAGTGAT
GGCAGAGGACTTTGATGTCAGTGATGGAAGCCACAAAAACTTCTAGCTTA
GCTCCATTCCCAACAAGATTGACTGCAAACCCCATGCTAAAACAACAGCA
AAAAGAAAGAATCCTCATTTCCAGGCATAAAATTTTTCCCCCAGTCTCTG
CTGTCCTCCATAAGATGTCTGATTTCAACAGGAATTACGAGGCTATAAGA
AAGGCAAGAAAAAACTACACACTGTCAAGAGAAAGCCATCAGAATAACCA
GACTCGTAGCACAGACACTGGAATTGTCAGGATATTTTAAATAACCGTGA
CAAATACATTAAAGATTCTAATGAGAAGGGGGTAGACATGTAAGATCACA
TAGATTTCAGCAAAGAGATGAAACTCGAAGGAAAATTAAATGGGAGCCCT
AGAGTGAAAAACACTGTAGCAGAGAAGATGGGTTCATCCGTAAACATGAC
ACAGCTTAGGAAAGAATCAGTGAACTTGAAGACAGGGCCACAGAAAATAT
CCAAACTGAAATGCAAGGAGGAAAAATAATGAAAGGGGGAGAGAGAAAAA
ATAAAAGAACAAAGCATCCAAGAGCTGGAGGGTGACACTGAAGAAGAGAG
CATAGGCATAGCTGGAATCTCAGAAAGAGAGAAAGAAATAACCCAAGATG
TAATGGATGAGAATTTCACAGAAGCGTTGTCAAGCAACAAACCATACATC
CAAGAAGCTCAGAGAACACCAAGCAAGGTAAGTACTGTAAAAAAATAGCC
CGAGGTATACCTCATTCAGGCTGCTGAAAATCCATGACAAAAGAAGTCTT
GAAAGTAGCCAGAAACAGAAGGCGTGTTCCATTCAGAGGGAAAAGACACC
ATTGTTGCCAGAAACCAAATAAACCAGGGCTGAAAGGGTAAAACTTTTTT
TTTTTTTTTTTTTTTTTGGCCATGCCTGTGCCATGTGGAGGTTTCCCGA
TCAGGGATCAAC

Contig 113 (1300 bp)

AAACGGATAAATACAGGTGACCCACAGGCAGAAGCTGAAGTACAAACAGT
TCACAACGGCACCCAAAAAATACCGAAGGCTCAAGGGTAAATCTGACCCC
AGATGAAAGGCCTTCTCACGGAAAATGGCAAAGTGGCGCTGAGAGGCATG
AGAGGTTCGAATAGATGGAGGGCTCCGCCGTTTTCCCGGGTCCGAGGATT
CAGTGACGTCACGACGCCAATTCCTCTGAAACGCCTCTCTAGGTTCAGTG
CAGCCCAGACCCACTGGCAGCCGCCCCTCGCTGCAGAGACAGCCCAGCTGG
GTCTTGAGGTTCCTACAGCGAAGCAAAGGGTCTAGAAAAAGCAGACGTCT
CTGGAAAGGGAGAAGCAGCCGATGGATTGGCATACGGCGACAGGAGATTC
CTCGGACAGTGGCACCAGGAGAGGGGTGGACAGAGACTGGTGCAACCGAG
CGGGCCCAGGAATAAGTCCACACCCACACGTACCATCTCGTTGTTTATTT
ATTTTTTCCTTTTCAGGGCCACTCCTGGGGCATGTGGAGGCTCCCCAGCC

FIGURE 6, CONTD.

AGGAGTCGAATCGGAGCTGCAGCTACAAGCCTACCCCACAGCCACAGCGA
CACAGGATCTGAGCCATGTCTGCAGCCTACACCACAGCTCCCGGCAATAT
TGGATCCTTAACCCACTGAGCAAGGCCAGGGACTGAACCCACGTGCTCAT
GGATACTAGTTGGGTTTGTTACCACTGAGTCACAGTGGGAACTCCTTTAA
TTTTAATTTTTGAAGGTTCAGAACTCTTTAATTTTTTAGTGAGGTATAGA
TTATATTACGCACCATTTCTTTCTGACTTCGGTGCACGGCTTTTCAACAA
ATGGGTGCTGGACCTGCTGGGTGCCTTCTTCAAATGAACCACAAGCCCTC
CCTCGCGCCGTATGCAAAATTTAACTCGAGGGGCTCATAGACATAAACGT
AAACTCTAAAGCTATAAAATTTCCAGAAGAAAACGTAAGGAAAACCTTTG
GGGTCTTGGGCAAAGATTTCTTACCCATGACAGCAAAATTACAATCTACA
GAAGAACTGGTGGCCTTTATCGGCATTTAAAACACCTGCCCTTTGAATGA
TGCTGTCGCAAAACCGAACATGCAGCAAAACGGATGCAACTAGCAGGTCT
CACACTCAGTGACCCACGTCAGAAAGGGAAAGACACGCCACGTGACATCC
CTTAGATGCAGAATGTAAAACACGGCCCCCGTGAACCGACCTCAAGAGAG
AGACAGACCTACAGACGCAGCAAATTTGGGGTTGCCGAGGGGGATGCCGG
Contig 114 (3000 bp)
TGTGAGACCCCTTGGCGGGCCAGGACCCCCCAAGGTGACCGAAGGCCTCA
GCGCCCCCAGCCGCCCCATCCCCCTCTTTCCCGACACAGGATTTTTTCC
CACCAAGCTCTGTTCCCTTGGTCACGCTCTCACTTGAGCAGCCTCAGGGT
CTCCCGGTGCCTGTATCCACGACAGCGTGACCTTCTTGGTGTGTCAACCC
AGGACCCCACGCTGGCCAGCCACGCCTTCCCAGAGCACCCCCGCCCATCC
TCAGAGTCCAGAGGAAAGGCCCCCATTGACCCCAGAAACCAAAACGCAGA
GACTCTGGGACGCCAGCAAGAACGTACACTGACTCCCACCTGCTTCAGGC
ACGGAGGCAGGGGTGGGTTATGAGCGACCCCGTGGAAGGGCCTTCTTGTC
CATCGAGGGGCTTCCAGGGGCTCCTAGACGGGGATGAGTGTGGCAACATG
TCGCCGCATTACAAAAGACCCTGCAGTGCTGCTGGGATGGGTCCCCCGGC
TAGAAAAGCAAAGGATTCCAGCCCAGTCGAGTAGGAGGCGGCCTCGGAGG
CTGCAGAGGCGCGGGGGGCGCTGACCACCACTCGGCAAGCCCCGTGTTGG
AGGGGACGCCCGGCCCGGCTGCAGCCGGTGCGCCTCCGGATAAGCTCCTA
AGAGGCCGCGTGCCCCATGCACGCGCGTGCACACACTCGCTGCCCGAGGG
TCCTTCAGCACAGACCTTGTGGGGACGGAGGACCTGGCAGGGGTGTGGCT
CTGGGGAAGGGGTCTGTCCCAGGAACCCTGTTCTGGATTTGGGGGTGGGC
GTGGATATCCCGTCCCAACCTACAGAAGGGAGGGGCTTAAAAAGAGCCCC
TTTGGTGTGAGGGGCCAGCAATCCTTTGGCTTTTTCTTGGCCCACTTGGA
GCTTGACGTCTGGTCAGTGACTGGGAGCCAGGGCCAGAGGGGGGCAGCCG
GGCTGAGGCAGGTTCAGGCCAACCATCTCTCGGCCACACTCCCGAGGTCG
GGCAGCTACGGGGCCCCCAGAGACACAAGCCCCAGGGGTCCTTCCCCCCC
GCCCCCTGCCCCAGATCACCAGGAGACCCAAGCAGCTCTGCCTCCCCGTG
CCTGAGAAATGCCCCATCTGGGTACCCAAATCACCCTCCCAGAAGGTAGA
GTGGGGGGCCCAGGACAGGGGGACCCCAGTTACAGAGCCCCAGGCAGGCT
TCCCAGGGGCGAGGGGACTCCGTTTGGGGCACAGACGGAGGCAGAGCGGG
CTGATGGATTCTCCCCCGGTTCAGGGATGCTGGCTGCCTGGCCTCCAGGA
GCCGGCGGTGCCATCTGATCTGATTAAGGCCTGCAGTCCCAGCTGGGCGG
GCACAGCCTGGGGGCTCGGCGGGCAGGGAAGAAGGCGCTGTCGCCCCCAGC
CGGTCAGGCTCGCTTTCTCTTCATTTCCTCTCCATTAAAAGTGTCAGAAC
CATTTATTGATTTTTTAAATCAGGACGTGCTGTCCGTGACACAGCAAAGT
GAACAAAATCAGAGCAAAGAGAGGCCAGGGCTGAAGCCCCAGAGGGCGGC
GCCTCCAATCCGGGTTGTGCCCCGGGGCTCCAAGCCCCTTCTTCTTCTGG
GGTCCTCGGCGTAGTGGCCAGGGCAGAATGCACCTGCCGTCATCCTGGGA
GGCTTGGCCATCGCTGGCTTCTGTCTCATGACGCACCGTCGTTCCATATC
TACGGAAACAGCTTCGCATTAACAGGCAGGGGAGGCGGTTGTTTCTCCTT
TATCTGCCCACCATCGGCGCTGGGGCCACGTGGAGCCCAGCCGGCTGACT
TCCCGCTCGCACGCAGGGCACTGATTGCAGGAACGAGGACATCCAGCCCC
CGCCTCTCAATGCCCCGGGTGCTGAGAGCATTTCGCCCAAACGGCTTGGG
TGGGACAAGGGATGGAGCTGTGCGCCAGGGGCCTGGCTGGGGCAGAAGGG
GGCCTGCCCGTGTCTGCCCGTGGCCTCCAGCACCCTCGGCTGCCAGGCTG
CTCTGGAGAGGTGCCCGGGGGCCGAGGGCCAGGGGCACCCTGTTCTGCCC
CACGTCTCTCTGTCCTGCTGAAAGTTCCACCAGACGCGTGCTATACCCTG
GGAGTCAGGAGGATGGGGGATAGTTGGGGCTTGACGTCTGTTTCTGAAAA
AACACCGTTTTCCCTGAAATATATATGTATTAATTTTTCGTCAAGATAAA
ACTGTGTATAGTTTTTCGTGATGAGAAAACGCATCCATCTTCCTTAGAAA
GCCTGAAGAGGTACAGGAGCCTATAAAGGACAAGATGACAGATGCCTCTA
ACGCACACCAAATGTGCGGTGGCCCCCAGGGGACCGCATAGACGGGGCGG
CTCCAGATGGCCACCGTGTGCGAGGGACACGGTTCAGGGTGGCAGAGTAT

FIGURE 6, CONTD.

TCCTGGGGGGGGGGGGCTCAGCGGTTCCCATTTCCCCCTCCCTTCCTTCC
TTCATTTCTTTCCTTCTTTCTTTCTTTTTGTGGTTTAGGGCCGCACCCG
CGGCGTGTGGAGGTTCCCAGCCTAGGGGTCTAATCAGAGCTACAGCTGCC
GGCCTCCACCACAGCTCACGGCAACGCCGGATCCTTAACCCACGGAGCGA
GACCAGGGATGGAACCTGGGACCTCATGGATCTTAGTTGGGTTTGTTCCC
GCTGAGCCACAACGGGAACTCCAGCCATTCCCATTTCTTGCTCCAGTTCC
AAGAATTCCAATTCTTATTCCTGTTCTTTAAGGCCAGAGGCGACAGCCAC
GCCGAGTCCCAGAAGCAGGGCTCAAGGATGCTGCTGTTGACTGTGTCCGT
GGGCGGGGGGAGTTGATAAGAACCCCCAACACAGGGTGGTGGCCAGCAAC
GGGGGAGGGAGGAGGGGGGCTGGTGGGGAAAAGTCCCCTGAACCCCATGG
GCTGCCCCCTCCAGGCTGGGGCACGACCCCGAGCCCCATGGCCCGAGGAG
AAACGGTCCCAGCCCCAGGCTGGGCTCCCGCACCCCTGCCCTGACCCCGC

Contig 115 (1895 bp)

TCATGGAAGCCCCTTATCACAACCTCGGATCCAAAACCCACTGCGCGAGTC
CAGGGATAGAACTCGCATCCCCACAGACCCTATGTTGGGGTCTTAACCAG
CTGAGCCACATGGAAACTGGGTAATCTATTTTTAGATGTTCCTAGGGTTT
TTGGCCTTGCCTGTACGTGGGGACGCTGCTGGGCCAGGGATCAAACCCGC
GCCACAGCTGTGACCCAAGCAGAGCAGTGACAGCACCGGATCCTTAAGCA
CGAGGCCAGCAGGGAGCCCCTGTGTTTAGATTTTGGTGAGGATACTGCGT
GGGATTCAGGATATTCACTTTGGGGCTGTTGGAATTGCCCGTCGCTGTTT
AAGCAAAGAGAAATCCCTTCACTCTGTGTAACTGTGGGGAAATCCTTTAG
TCTCTTGAAACCATTGCCTGTGTTTAAGAGTGGTAACTCTGCCACCATAA
ATGCCCAGACCAGCGCCTTCCTGAGATCCGCTTTTGTTGCAAATATCTGG
TTTGAATGCTTTGATCGCCCCGCACCAGACCAGGGTGGGCGGACGCCGCCG
GGGACCCGACGTGACCATCGTGCTTCTGTATCCGCCCTTTCTCCGGCACG
CGCCCCCTGGTTGCCTCTGGCTGCTTTTAGTGGAGGAACTGAAGCCTCGC
CACCCAGACCCCGAGACCGCAGGACCCACAATGCTTCAAACACCTGCCCT
CTGACTTTTACAGGTCAAGTTCGCCAACGCCGAATTTGCACCGATTGGCT
ACAGAGAGCACGGTGGCGCCAAGCCTCCACTTGGAGTTTTATAAGGTCTC
CCTCCAGCTCGCAATGAAAATGAGCTGTGATAAGGCAAAGACAAAATTAG
TATGAAATCCAGATGCTTCATCTACAATACAATGACCGCGGGATTTGGGT
CTGAGCGACTGAAATCAAGGTGGGCTTCCGGAGGGAGGCTGTTAGAGGAA
AGGCATTCACGGAGGCTCAGGTCCGAGAGGCTTCCACACCCCTAAGAGGG
CTGAGACGGCAAGTAGGGACCAAGCCCCGCAGTCGGGAGAGCTGGGCAGG
AAGGAAGTCTGAGGTCACCCCCACCTGGGGAGGAACTGCCTAGAGAAGCG
GGGGCGGGAAGCAGGGGATGCCCAGTCCCAAGACAGGGACAGGGCGGAAA
GGGCTCTCTGCAGGCCCTCAATGCTGCCACAGTGTCCTCGTAAGAGGGAG
GCAGAGAGAATTGACACCGGGGAGACCACGGGACCACGGAGGTGGAGACC
GGGCTGCCCGCGCGTGCCAGTTGCTCCCGAAGCCGGCCCCTCCCCCAGAG
CCTTTGGGAAGAGGCGCCAACCTGCAGTTCTGCTACTCGGGGACAGGGAC
AGGGACAGCCCCCTGGAGCCGCCTCTTAGGGGCAGCATCCCCCAGAACCT
TCCTTAACAGACCATCTGGAGAGAGATGGGTCTGGGCTGCAGCTCCTGGA
ACTGTTTTGCCCACCCGGCGAGCACCAGTGGGTGCCAGCCTGGGCTGCCC
AGCCTCAGGGCCGGGGAGGGCTGAGGGCACTGGGGCCCGGCTCTGGGACT
CCCCTGCCTCCTGCCCGTGCAGGACAGCCACCTCCCAGCATCTGCTTCCT
GCCACCCACATCCCCAGGACCGTCAGCCCAGGCATGCCCCTGGCGTCGGC
CACTCACACCACAGGCCAGGAACCCAAGGGGGCAACACAGAAGGGCAGTT
GCCATCTGCAGATGCAATGGACAAACTGGGGTCCGTGATGATGGCAGGCT
CTGGGCGCCCGGGCTGGCAGGGGAGCCAGGACTGTGCGGCCATCACAGGA
AGGGCATGACGGGGTGAAAGCAAGAGTGGAAACCTCTGCCACCCGCCTGG
GCGCACATACCGGCCACCCTGCAGCCCCACCCCCATTTGTTTGCT

FIGURE 7

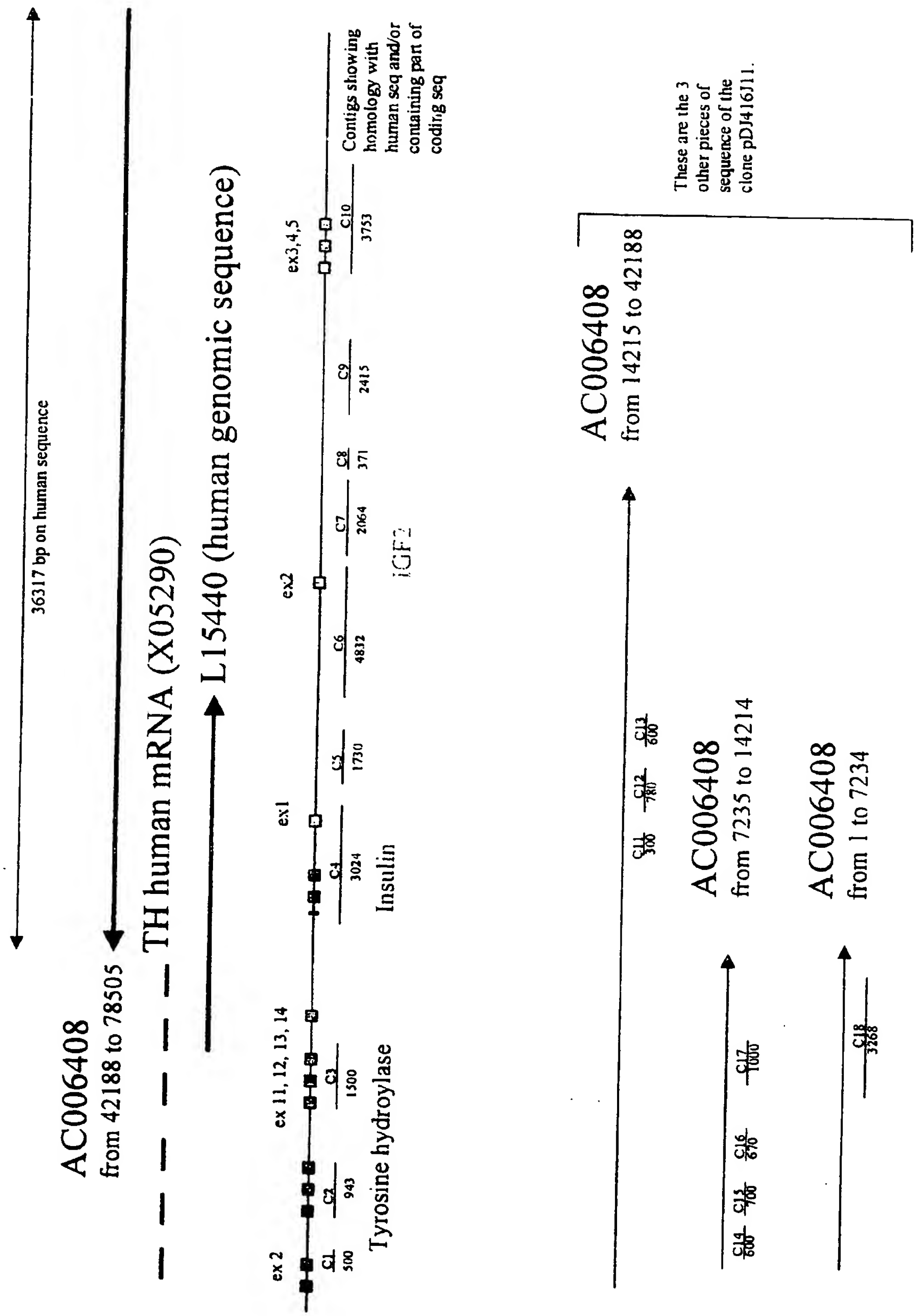
36317 bp on human sequence
AC006408
from 42188 to 78505
TH human mRNA (X05290)
L15440 (human genomic sequence)
ex 2
ex 11, 12, 13, 14
ex1
ex2
ex3,4,5
C1
500
C2
943
C3
1500
C4
3024
C5
1730
C6
4832
C7
2064
C8
371
C9
2415
C10
3753
Contigs showing homology with human seq and/or containing part of coding seq
Tyrosine hydroylase
Insulin
IGF2
AC006408
from 14215 to 42188
C11
300
C12
780
C13
600
AC006408
from 7235 to 14214
AC006408
from 1 to 7234
C14
600
C15
700
C16
670
C17
1000
C18
3268
These are the 3 other pieces of sequence of the clone pDJ416J11.

FIGURE 8

Contig 1 (1040 bp)

GCGCGCCGGATCCTTAATTAAGTCTGAGAGATCTGCGGCCGCGGCCAGGGTCTGCTTCTG
GCCAAGTGTGGGGCTCTGCTCCATCCTGGCTCGGAGGTCCACCCATGGCAAAGCCTGGGG
TCCTCCCACTGAATATTTGGGGGTCCACTCGTGCCAAAGGCTGGGTGTCCAGTGTGCCAA
CGGTACATGGAAGCAATGTCTTCCCAAGGACCGTCCAAGGTGTGGTCAGGCCTGGACAGC
TGTGAGTCCCTTCGGGACTAGACTTGGTGGCCGAACCCTAGGGACCGTGCCCGAGGCCCC
CCACGAGGCCAGGTGTTTGCCCCAGGGACAGAACGGCCAAGGGTGGCCGAGGGTTCTTTT
TGTTTGTTTTTTCTTCTTTCTCTTTTTCTTTGGCCGAGGGTTCTTAAAGCGCTCTCTCTG
CTCTTTGTCCCGATCCTGAGCGGGCAGTGTCCTGGTCGGTGGGGTGCTGGGCAGCCGCAG
CAGGGCTGAGAGAGCCCGGCTTGTCACTAGGGCGCGCCGGTGAGCCCAGCGGGCATGCCG
TGTCCAGACGTTGGATGGGGCAGCGAGGGGACTGGGGTGCCCCAGCCCCCGTGGGAAGCC
CGCCCTGTGGAAGCCGCTGTGCTCGCCACAACAAGCACCGTCGACTAGCTGGTGAATCAG
CGCCCGTCGCCCGCGTAATCCCAGGCGCTTTCTGCCCAACCTGAGCCCTGACCCCACACC
CCTTGCGACCGCTCCGTGGACCCTGGGGCGATGAGGTGAACCGTGGGCTTGGCCATCGTG
GTGGCAGACGGTGGCACACCCGTGCGCCTGTCGGCCCCCCTCCATCCAGGAGCAGAGTGC
GCACCCAGTGGGGGCTGGGCAGGGAGCCGCCTCCACCTCCGCCCTGAGGGGACGGGACTC
TTTCGACCCGGAGTGGGAAGGGACATATGCGGACGATGCCAGACCCTGTCTGTGGGGGGA
GGGGGAGAAGGCCCTCTTTGGAGAATTCCAGGACGGGTGAGGAACGTGTGCTGGACCGGC
CGGGTCGGAGGTGGGCCTTG

Contig 2 (9234 bp)

GGCAACCAGGGGAAGATGGGGAAGCGGGGTGCAGGGGCGTTTGCGCGGGCCAAGGACCAC
CTTGGAAATCTGGAGCCTGGCAGGAGCGGCGCAGGGTTGAGGGGCTGGCTTGGGCAGGGC
TGGCTGGCACCTGGGAGCCTGGCGGGGTTGAGGTCCGGGCTCCCAGGTGCCCTATAGGCA
GGGCAACATCGGCATGGGGGGTGACAGGCCCGAGCTGGGGTGCGGAGGGAAGAGGGGGGA
GCCAGGCATTCATCCCGGTCAATTTTGGTTTCAGGTCGTGGCGGCTGGTGGTCAGGGGGA
GTTGGAGAGAGGTTCGCCCCGGGGCCTGGGGCAGCGGAGGTGTAGCTGGCAGCTGTGGGC
AGGTGAGGACAGCCGTCTGCCGGGCCAGGTGAGTCCCCTTCCCTCCCCAGGCCTTGTTTC
TCTGGCCTCCTGCATCCGGAGGTTCTGGGGAGCGAGGGCCGGCGAGGCGAAGCGGCTGAC
CCCCCGGCAGAGTGGCGGCGGACGACAGGCAAGGCGGGCAGAACAGGTGACACGTCTCAG
GGGGAGCTGGGACCGGGCGGGGCTGGGGGGCCGGGGCCGTCCCAGGTGGAAAGAGCATCT
CAAGCGAGTCTGGTGGGAGACGAGGCAGGGCTGCCAGCAGGGAGGAGACGCAACAGGCGG
GGGGCATTCCAGGCCCGGGTCGGACAGGACCCGTCGGGGGTGTCAGGACAGTGGGGTCCC
CAGCCGCCACTTCACCCACTGCAATTCATTTAGTAGCAGGTACAGGAGCGGCTCTGGCCG
GGCCTCTTGAGGCCTGAGCTGGAGCCTCGAGGGCCGGAGAATGGGAAAGAAGGTGCAGTG
TGCCAGACAGACGTCACCTGGAGGGAGCACGGCCGTGGGGACGGGCCCCAGAGAGATTTC
GGCAGCAGGGAGGCTGCGCGGGCCCAGCCTGCGGACGTGCGTTCCCACGCAGCACTGCGG
CCCAGGGGCTGGCGCGGCAGGGCCCCCGGTGTCCTTGGTGGCACTGTGCGCCCTCGCCGC
TCGCCCCTGGGACTGGCACGGCAGACAGGACAGCACCCAGGGGAGTCAAGGGCACTGACG
AGACCAGACTAGGCGAGGCGGGTGGGGTGGAATGGATGTGACCTCTGGGGGGAGGGAAGGT
GGGGACGCAGGCAGGGGCGAGGCGCCGGAGCCTGGCGGCGAGCGAGGCCAAGGCGGGCCT
CTGCGGGTGACAACTGAGCACATATGGGTACCTTTGCGCTCGCACCGGAGACAGGTGAGT
GTCTGGCCCCGGCCTGCCGCCCTCCCGGCCCCGCCACTGCCTCTGCCCTCCCCCTCGACC
AGGGCCCTCTGCTTCCCCACAGCCTCGTCTCCAGTGGGGGTGGACACACTGCCAGCACCA
CAGGCCGGACGCCAGGATGTGCTTGGAG**GGACATGACACAGTCCGGTGTGACGGAGAGGG**
ACAGACGTGACGCCGTCCGGCCTTCCTGGTGAGCGCAGGTCCAGGCCTTGGCCCCCAGGC
CAGCCGCCCCACCCCCCACCCCTCATGGCCGTCTTCTGTCCCGCAGAACACTCTCGGCTG
GCCCCGCGGGGGAGCTGCCACACCCAGCGTCTGTTCCTTTGCCTTCCTGAAGGAGCACGT
GCATGACTGCTGCTCTCTGGACCCCAGAACCCTCAAACGACAAGGTGAGGCAGGTCCCGC
CTCGCCCCACACGTGGAAGGGGCGTGGGCGAGAGCCGGGCGCTCACGGTGCCCCCCTCCC
CCTGC**AGAGATGGTGCTACCCAGCTCATGCCTGGGCCTT**GGACCCGGACTTCTTCAAGTC
CTCCTAGCTCTGACTCAA**GAATATGCTGCATTCTGGAGCCACTACACTACTTGACTCAGG**

FIGURE 8, CONTD.

AATCAGCTCTGGAAGGTGGGCGCGCGCTCCTCCCGCTCCCGGAGCCCCGCCCGCTGCCCG
CTCCCCGCTCACGTCCTGTCTCTGTCCTCGTCCGCAGGTTGAGCCAAAGGAACAGACGTC
CCACACCACCGGACCAACGGCACCCGCGGGTTCCCCACCCCCCGCCCGGCCACTCCACC
TCGGCGGCCACCCCCTGCTGCGCCCTGGAGACACCACCAGCCTCCCTCTCTCCCCTTCCT
CCTTTTTTTCCTCTGTCTTTTCTCTTCTCTTCTTTCCTCTCCTTTGCTCAGAAGACTCGG
GGCA**TCCAGGACTCTGTGTCCCCGTCCTTCCTGAATTAATTTGCACTAAGTCGTTTGCAC**
TGGTTTGGAGTCCTGGAACCAGCCCCGGGTCTCGGAGCGGGTGTGTGAGCTGCCGAGTGG
CCTGGCCTCCTCGGCCCGCGCCCCCTCAGCACCTGCCATTGTCCATCTCTGTCTGGGGGT
GACTGGGTGGGGGCCTGAGTGTGTGGGG**CCCCGCCCTCCCCTCTCCTAGTCTGGAAGCTC**
CGACCACCGAGCAGACCTCAAACGCTGCACTGAGTGTCCATCTCGTCATGTGCCCCTCCT
CGCCAGGGCCACCCCAGAGCCCTGGACTCATCAATAAACTCAGTTACCGGAATCTGTCTC
AGGGGCTTTGCAATTGGGCTGGGGTGCGCCGGGGAAGGGGGGATGAGATGGGGAACAT
GCAAGGAAGGGCCTGTGGGCTGGGGACACAGAATGGGTGGGAGCGGGCTCACAGGACT
CGGGGGGTAATGAACGTGGGGCTGGGCGCAAAGGGGAGTGGGACGTGGGGATCAGGGCGG
GGGGCCTGGAGGATGCAGGGTCCCTGCAGGGAAAGGGGGCCGAGGGCGTGAGGCATGTCC
TCAGCCCTGAGAGGCCCTACCCCACAAAGCACAGCCTGCGCGCCACCTCCAGGCCCCCAA
ACCCCCGCCCCAGACCCTGAAGCCCTGGTCCAGGGCAGTGGGTCTGACTGGCGGAAGGAA
CATGCCACCCAGGCTGGCCACACCACTGGGACGCCCATGGGCGGCCACTTTCATCAAGAG
CCTGGCAGGCCCTGAGTGCTGGGCTGGAGGGCACAGAGGGTCCCCCTCCCCTCACGCTTT
GCGGTGCTGGGGCACCGCAGGAGTGCCCAACAGGAGACCCCAGGAAGTCTGCTGGGCTGC
AGCGAAGGGCAGGGTAGGGGGGCGGCCCACAGGGGCCCAGCTCAGTAGGCAGGTGGCAGT
GGGAGGCGGCAGAAAGTTGGAAAGGGTGGACTGGGCACGTCAGGATCTCGTGGCGGCAGC
CCCGGAGCCACGGCCTTGGGTGCACTGCAGCCCCCACGGTTGGTGTCCCGGTCCCAGGCA
GCAGCTGGGCTGGTGACGCCCCTCTGCCTCTGCCCACCCCCCCCACCGCCCCCCCGCCAG
CCTCCCAGCCCCTGGGCGCCTGGCGTGACGCTGGGAACGCGAGGGAGCAGGCCTCGGAAA
CAGGGCTGGGTCCTTGACCCCTTCCTCTGCTCAGGGCAGTCAGGAAATGCCTAGCGGGCC
GACTGACCGAGAGGAGATAGCGGAGGCCTGGGAGACCCCGCGCTCGTGCCGTTCCCAGCG
TCCGGCCGCGTGGCCCTTGGCTGGCCTGGTTTGGGCCCCATGAGCTCACCCCCCGCCCCC
CACAGCCTCCCCGCGTCTGGTCTCCTCTCTGGGCCCTGCTGTCCCTCCTGACGGGGGACA
GAGCCCTCCAGGGCCCCGGGGGGACGGTCCCGGGTCAGCAGGGCGGGTGGGCAGCACAGC
TGCGTTTGGTGAAGCCCCTGCCCAAAGCACCCTCAGCGTTTCCTCTGCGCGTCCGGCCGC
CCCCGGAGGCTTTCCCAAGTCCACGGGCAACTCGCAGGCGAGCCCACTCCACCTCCATCA
CGCGGGTTTGGCCAGCGGCAGAAGCACTCGCCCTTCAGGCGTCAGGAGTTAAGCCCCTCC
AAGGCCCGGTGCTAATCAGCTGCCTCTCCTGGAGCTTCGCAAAGCGGGCTCTCAGAGCCC
AGCTTCCCGGGGGCTCACCGTGGTGGCATGGGCACCACAGGTGGCCGGAGGGGCACCGAG
CACGACGGGGCTGTGGGGGGTGGAGGAGGGACGTTGGTGACTCCGAACCTCTACTGAGGC
ACACAGAGGACACGGCCGCTTCCAGGGGAGTCAGCCTGCGAAGGGCAGAGGGGCTGTAGC
CTCCCGGTCACGCCCTCGCCTCTGCCCTGGATTCCTCCTGGGGGCCCGCGGCTCGTCGGG
GAGGTGAGTGCCCCTGGATGGGCGTAGGCTGGGGGGGCAGGGACTGGGGGAGCCCCGAGG
CCCTGGGCCCACAGCCCTGTCTTGCCCCACACACAGGGCTGTCTACACTGGGTGCCCACT
TGCTCTGCTTCTAGGCTGTTCCCTGGGCAGCTGCCTGGAGGGCCGTGGGCACAGTGCGGG
CAGCCAGTGGGGAGGCCGGGGATGGGGCCGGGGATAGGGACCCCTGCCCCTGGGTGAGCC
CCACCTGGGCTGGGAAGACAGCAGCAGCGCCCCTTCAGGTCCATGGACCAGGGGACCCAG
GGTGGACTGTGTTTACCTTCAGCCCAGGCCAGTTTCCTGCTTGAGAAAGCCCGGGAGGGG
GTGCGGGACAGGCCCGGGCCCCCCACGCAAAGGCAGTTTCGCAATGTCCCTGCGCTGACT
GAAATGTCACCAGGCACACGGCTTGAATTTCTCCCCCAGACCTGGCAGGGGCGGGGGTGG
GGGCACCGGGCTGCTGGGATCTTGGCCCCTGAACCTCCCCCGGCCCTGCGGCCAGGGAGG
GTTTAGGCTGAGTGACAGCCCACGGAAACCTGGACCCGACATGTCTGTGTGTCCATGTGT
GTCTGTGTGTGCGTCCACCTATGCGTCTGCGTGTGTGTCCATGTGTGTCCACATATCTGT
GTCCACGTGTCTGTGTCCACGTGTCTGTGTCCACGTGTGTGTCCACGTCTGTCCATGTGT
CTATGAGTCCTTGTGTGCATCTGTGTGCCCGTGTGTCTGTGTGTCTGTCCCCTGCAGTCC
CCGTGGACCTGTCTCTTATACACATCTCAACCTG
GCAGCGCCCCTTCAGGTCCATGGACCAGGGGACCCAGGGTGGACTGTGTTTACCTTCAGC
CCAGGCCAGTTTCCTGCTTGAGAAAGCCCGGGAGGGGGTGCGGGACAGGCCCGGGCCCCC
CACGCAAAGGCAGTTTCGCAATGTCCCTGCGCTGACTGAAATGTCACCAGGCACACGGCT
TGAATTTCTCCCCCAGACCTGGCAGGGGCGGGGGTGGGGGCACCGGGCTGCTGGGATCTT
GGCCCCTGAACCTCCCCCGGCCCTGCGGCCAGGGAGGGTTTAGGCTGAGTGACAGCCCAC
GGAAACCTGGACCCGACATGTCTGTGTGTCCATGTGTGTCTGTGTGTGCGTCCACCTATG
CGTCTGCGTGTGTGTCCATGTGTGTCCACATATCTGTGTCCACGTGTCTGTGTCCACGTG
TCTGTGTCCACGTGTGTGTCCACGTGTGTCCATGTGTCTATGAGTCCTTGTGTGCATCTG
TGTGCCCGTGTGTCTGTGTGTCTGTCCCCTGCAGTCCCCGTGGACCTGTGTGGTCTCTGG
TGTGCAGCCCTAGCCGCGGCCCGTCCCAGGCTGAGTGTCCCCAGGGTGCAGCACAGCTGT
GACGAGGGTGTGGGTCCCGCTGGCCGTGTCGCTGGGCTGTGGGCCCTATCCTCTTTGTGG
CTGCTCTGCAAGGCCTGATGGCTTTTGTGTGGCCTGGCCGTTCGGGTCCATGCCCCCTGG

FIGURE 8, CONTD.

AAGAGCAACGTCTGAGCTAGCTCCACGCGTGGGTCCATCTCGGCCCAGGTTTAATGAGCC
ACTTTCAGGCAGGGATTGCACAGGAGGCAGGGTGGGAAGTGGCTCTGCTCAGACCCCTGA
ACAGGGTCTGGAGATTCTCCAAGGCACAAAAGAACGGACGATGCCCCTGGGGTCAGCGA
CAATGCTCCCTGAGAAATCTTGGCACACAGGGCTGGGCCTGCGAGGTGGCCCCTCGCCCC
ACCCCAGCCTCCTGGAGGACAACCGTCGCCCTGCTCCCAGAGCTGGGGGGCGCCACACGT
GGGGCACAGGGAGCATGGCCCGATTCCAGGCCTGGGCTCCCTCTCGTGTCCAGGATCTC
CCCGTGTCTTGTCTCAACAAGCCCCTGACTTGGAGGCCCCAGGGTGACCCCTTAAAGGGG
GAACAGAAGGTTCTAGAAGGAGCGTGGCCAGCTTTGCCTTCCCTAGGGCTGTGGTGACCA
CACTGGGCCACGGCCCAGGCCACCCCACCCGCCTCCTTCCCCCTGGCCCCCTCCCTTCCC
CGCACCTCTCCCTGGCCTGCACCTGGTGACACGGCTGGCTCCCAGCCAGGGCTGAGGGGG
ACCAGCGGGGCCCCTTCCTGGAAGCCCACCTGCAGGCCGGCTTGCTGGGAAGGGGCCTGC
TCCTCGCCGGCCCCACCCGCCCGGGGCCGTTTCCTGGAAGCGGTCACTGGATATTTTGTT
CCTTGTCAGCGCCGAGCTTGCATAAAGCAGACACTGAGCTCCTTGTCCTCCGGGAGCACG
CGCTCCATCACCGAACACCTGGCCGGACACAGGCGGGCAGCCGGGCCTGGGGGAGCAGCG
CGGGCCTGGGGCCGGACCAGCAAACGATCACGGCGCCGAGCGCAGGCCCCGCGCCGCTTC
TGCAGGCCGCCCCCACGTGCCCAGGCCCAGCGGTGCCCATCCTGCAGGCTGGGAGGAGGC
TGTGGGCGCAGAGCTGAGAAGGGGGCAGAGGCACTGGGGGGGACAGCCGTGTTCCCACA
CTTTGCAGAAACCTTGGCCGGCCTGGATGTCTTGCTGGGAGAGCTGGGGGAGGGGACAGG
GCAGGAAGCCGGTCCCCCCGAGCGGGGTAGGAAGAGGCCTCGGCCCTGGGAGGAGGAGGA
GGGGAGGGCAGTGAGATGGAAAGAGCACCAGGGGCTCGAGGCTTCTTTCTGGAACAAGGA
CTAGAAGGAGGAGGCCGGGCAGCTGCTTGGGATGCTTGGAACAGGCCGGCCCCAGTGCTG
ACAGGGACGTGACCTGGGGGCCGGTCCCGGGCCCAGGCGGGCTGGGAGGGCGCCTGGTGG
GTCAGCGCCACTCAGAGCCCTGGCAGCAGGGGGCCTGGGCACGGCTGCAGGACAGAGCTC
AGGACACAGATGGGGGCGAGGACTGAGTGGGGCACCACAGATGCTCCCAGGAGGTGGCCA
AGGAGTGGCCTTGGGATCCCAGGATGGCCCTGGTCCCAGAAGATGCGGCAGCCCAAGGGA
CCAGGCCAGGGCCGCAGGGGGCCACAATCTGAGCAGGGCTCAGGCCCAGGGCAGAGGCCC
CCTCCCACCCAGCCCTCCCTGGGCCCGCCTCTCC
GTGCAGGCAGTGGGCTCAGATGGGGCAGACATGAGACCAGGTCCAGGGAGAAGCGGGGCC
CCTTGGCTTCATTCAGGTGGCTTTCAGACCGCGCCCCGTGCGTGGCAAGGCCCACAGCGC
TCAGGAGCACACAGACCCCCACCACGGGCTCCCCCAGGTTGGGCGGTGACATCAGCCCTG
TGTCAACAGCAGGAGCTGGCAGCTCCCCACCGGGGCTTAGGGAGCGGGGACCCTGAGCCA
CCCTGCCACCGCCCCACCCCACCGTGGCCCACACGAGGGCCCGCTGCTCTGGGTCTGGGG
CCAAGGCCCCCCAGGCGCCTGGCACTGTCTGCCCCTCCCGCTGGCTCTCCGTCTCCAGTG
TCCCCGCCAGAGAGCATGGGGCCACAGGCCTGAATGCCACCCTCTTCCTCCCTCTGGAGG
GGGCCTGAGGTTTTGGGGGTTCACAGAGTGGCCTCCGGGGTGGGTCCAGGCCCAGCGAGG
CAAAGCGGACCCCAGGGAGTCCCGCGGAATGTGGACAGCCCCCCCGTAGATCTCGGGGG
GGCCAAGCTCTGGTTGACCTCCATCCTGGGGCTGTGGGCCTTTGGTCAGTGGGGAGGGTC
ATGACACCCAGCCCACCAGCTGGTGACAGCCCTGGACGTGCCGGCTCAGGGCTGGCCTGC
CCCTGCAGCCTTGAACCCCTGTTCTCTGGGAGTGGGGGCGCAGGGGGCGCCGGGGCAGGG
TGAGAGACGAGAGCCTCTCTTCCCAGAACTTCTGCCTGCGATGAGGACCCAGCAGGGGCC
TCTCCTCACCAGAGGGCCTCTGCCGGCTGCAGGGCCCCAGAGAGGCCCAGAGGCTGGAGG
CCGGGCCTTGGGAAGAGGCCGGACTTCCAGAAACCAGCTGCCCGCTCCGCAGCACCCAGC
GCCCACTTGGGAGGGGGGCGCGCCCCCGTGCCCCGCCCGGGTCCACTGCTGGGGCCGCCA
CAATAAAGTTTGTCCCTGCTGGTTACTGTCCGTGTCTGAGAGGTTTCTGGAGCCTGGCCA
CAATGGGCGTCAGGATGCGGCTGGGAGGGAGCCTCGCGAGTCAGAGTGTGCTGGTCTCGG
ACAGGCCCCGGCGCCCCCAGCCCGTGCTCTGTGGACAGATGGGTGGGTGGGTGGGTGTCG
GAGGGGGTTGGAGAGGGTGGGCGGGACGAGGGGCTTCCTGCACTCTGTCCCAGGGAAGCG
GGGACCAAGGAGGGGACAGCCCCCGGTCACCAGGAGGGTCCTGTCCCTCTCACCCCCCGG
GACAGGTGAGCTCCCCGGAGCCGCCCTTCTGGGACAGGACCCCACGGCCAGGCCACGGCC
CCCCCCACCCCGTGGTCCCTCCGTCCCACGGCCGGCCTGGGGGGCCACGGGCCCAGGGCC
CCCGCTCCCCGTTGGCCCTCCGAGGGTGAACGACCTCGCCTGGGACGTGGGGCAGAGGGC
AGGCGCCAAGAGTGACCCCCTGGGACACGTGGCTGTTTGCAGTTCTGGAGGCAGCCGAGA
TAAAGCGGCTGTTTTCCCAGTGGGCTCAGGGCCAGAGGGGGGCGAGGGGCAGCCCCAGTC
AAGGCCGGGCCGCTGCCTCGGGCTCCCCTCTGTGCGGAGGGAGGGGGCCGGTTGCACAGC
AGCCCCTGCCCGCCGCCCGCCCGCCGGCGCAGGCACCGTGGGACCCGGCCTGGTGCCCCT
CCCCCGCCCCTGCTCAGGGGCCAGCCCTCTCTGGTTCCCAGGACGCCCCCGCCCCGCAGG
CGGCCAGAGAGTCCCAGAGTGTTAGCCTCCCACGTGTGGGATCCTGTCATATGCGACAGC
TTAACTCAGGCCGAATTTCATGGGTCCTGGATTTGGGTGGGCACGGCCCCTGCACAGCGG
GGCTGGAAGCCTAAGGCGGTGGGCGTGGGGGTGAGAGGCCCGCAGACAACAGGAGGGAGG
CTGGGACACTTCAAGGGTTGACATGCTATGCCTGTCACGGATAAATGC

Contig 3 (5347 bp)

AGATGTGTATAAGAGACAGGGGCTGGGTGGGAAGGACAGAGGGTGGGGCCGGAGGAAATG

FIGURE 8, CONTD.

GGATGCAGAGCCCACCGTGCACGCTCTGCTGGCCTTTGAGCCTCGCTGAGTCGCAAGAAG
CCCTCGGGCCTGGAAACAGACCCCCGGCCCCCACCCCCACCCCGGCCCCCGGATTACCCC
GGCATGGCTGGAGGGCCCGAGAAGCCACCCAGGCTTCCCGTGCCGAGCTGGGTGCTGGGC
CCAGCCGAGCGGGCTTGACGCCACGCTTAGCCCTCCCCAGGGAGCCCAGGGTCGGAAGGA
AGAGGCCGGCCGGAGGGCCCGTGGCCGCTCAGGCTGGAGGGGGCCCCCGGGTCAGGATGGG
CCCCAGACGTCCCCGCTCCCCGGCCATCCGTCACGGAGCTGTCACCCAGGAACGTGCTCC
AGACGTGCTTTCCTGCCGCCGAGGCCCCGAGCAGGCTCCAGGCGCCCCCACCCCCGAACG
CCCACGCACACCCTCGGTCTGCGAACACCCTGCCGTCATCCGGTGGCCCCGGTTCCCGCC
GCCCGCGCCATCCGGGTGCCCCTTCCTCCCTGGGTCGGGGGCCATGCCCTCAGCGGGCAC
GCAGGCCTGTGCAGGTCTGTTCTGACTCTTCCCCAAAGACGCAGGCCGGCTGCGGGCGCC
CCGACCTCGTCTGAGGCCCGTTTGTGCTCACTGGCTGTCTCAGAAAGGGGTGCCCACGGG
AAGCGCGTGTTCCTTGGGCCGCAAGGCAAGGGAGCCCACCCCAAGGTGGCTGAGGGCAAA
TGGCCCAGGGCCTCTAAGGAGTCCCTGGGGGCCGGGCCGGCCTGCAGCTTGAGGAGGAGA
GCCCTGGCTCTGCTCCCCCGGGCAGGTGAGCCCACGGCAGGGGGCTCCCCAGCAGCCTTG
GCAGGAAGCAGTGAGGAAGGGGTGAGGATGAAGGCAAGGGGGCCTGCGGGGACTTGGGCA
AAGCCCCTGAAGAACTGAGTTCCTCGGAAAGGCCGGAGCCCTCAGCCGAGCCTCGGCCTC
CGAGCGATGGAGGCGGCCCACCTGCGGCCCCAGGGTGCAGCTGTGCATCCGTCCCCCTCG
GGCCTCCCCCTGCCCCCCCGGCCACCACACTCTCCCCCTTTTGCCTTTGATCACTTGAGT
GCGACAGCTTGTGCGGCCTGAGCCCCAGAGACCGCTGCCCCCCTGCCGCCAGCCCCACGG
GAGCGTCCACCTGGGCCTGGCCTGGGCACTCATCCCTCCCGGATGAGGCCTTTCTAGCCT
GGGCCGCCCCGGGAGCGGCAGACCCAGCCCCTCGCCCCCCTCCCCCAGTGAAGGTGCTGC
CTGGTGGTCTGGGGAAGCCCCTGGAACAGGGGGCGCAGGTCCCACACGGGTGCTCTGGCC
TCCAGCTGCCAGGGAGGGCCGCGCTCAGGCCAGGGTCCCCTCCACCAGAACCGCCAGGGC
CCTGGGGAAAACCTGTCTGTGCTAACAGGGCCGCTCCCCGGGACTCCACGGAGAGGTGCG
AGGGACCCCTGAGCACCCACCGCCACTAAGGGGCCCAGCCAGCTCGCGGGTGCAGGCAGC
CGGCTGGGCGCTCACATGCATACTGCTCTCTGGCTTTGTGTGTGCGCCTGGGTTGGGGTG
AGCGGAGGTGCCCGAAGGCGGAAGAGCCCACCCTCCACTCGGGGACCTATTTCAGCAAGA
AGACGGATGGGACTGCCGGGCATGGACAAAGGAACAGGATGAACCTTCTGGAACGCACAA
GGCTTCCACGGCTGACCGCTCATAGGAAGGCGCGTCTCTAGGCCAATCCACCGTCCACCG
TCCATTCCCCAGCCCTCGAGAGGGGGCAGGATGGACCGCTGCAGCGTGAGAGAGCTCTGG
GGCGCTCCCACAGGGCAAAGTCCCAGGGCACTGACCTCAGAGCCCAACCAGGCCACCGGG
GCTGGGCCCACCAGGGAGCCGGGGCCAGGGTCAGGGTCAGGGCCCAGAGTGCGGGAAAGG
GTGCCGTGTTGCTTGGGGCGGCGGGCGCGCAGACGGCCCCTCGCACCCCCCGACAGCCCT
GGAGCTGAGTGAAGCCCGCGGGTCACCTTGGCTGGGGTTGGGGTCTCCTGCGACCGGCAC
CCCAGCTCAGGTCATCCTTGCTGTACCGCAGAGCGGCAGGGGTTCTGAGCAGGGACAGGG
TGGGCCGCGCAGGAAGCCCCCTTCTCTCTGAGGCTGCCCCGGCCCTGGAGCCTCTCTGGG
GCATGCCACCCCTCTCACAGACGCCTCCCAGGAGCCCCCACTTTCCTGCTGCGTGGTGAG
GGTGTCTCTCACCCGATTCCTGGCCCCTGCAGGTCGAGTGAGTCCCTGCTAAGCCTGGGG
TTGGAGCAGGTGCAGGGCATCACCACACAGCAGCAGAGGCTGTGGGGCCCCCTGAGAGGC
GCTCCCAGGTACCCTCCTCAGGGGGCTGAGCCCGGGGTTGACCCGGGACCTCGCCTGCCC
CAAAGCCGGCGCCCTCCTCCCGCCCGCCCGACCAGGGCCAGAGAAGCAGGTGTGGGGCGG
CACAAACCCAAGTCAGCTTCCAGATCCTGCTGGGGCCGCGTTGAAACTCGAAGCCCCCAG
GCTGGGAGGTCTAGACACCCCTGCCCAGACCGACAGCCTGGGCCTGGCTCACAGCTGCCT
GGGGGCCCAGGGGTGCACCTGCCCTGTGGGTGGGGTCAGAGGGCAGGGAACCCTCGGGA
AGGTCCCCCAGGGTCAAGGTTGGGCCTAAGCTCCGGTGACCTCTGGGAAGTCTGGGGCTG
GGTTTTGTTCCCAGAGGAGAGAGGGCCAGTAGCCTCAGAGGGGCTGTGGCACGGTGGGAA
GGCCCCAGGTGACCCCAGAGCGTGCCAAGCAAGCCCCCTTGACTGCAAAGC
GCAAAGGGCAGAGGTGGGGTGGGAGCCTCGACCCCCCGAGCCCAGGTACACAGGGGAAG
GGCGAGGGATCCGGCAGGGGCCACACCCGCCACCCCAGGCAGCCCACAAAGCCTTTGGGC
CCGGAGCCCCAGATGGGCCCAGCCCAGCTCTGGGAACAGTCTTCCCAGAATTCCCCAGCT
CTGGGTACCAACAGGGCTGCCCGGCCCCCAGAGCCCTCGGGCGGGAGACCCTTCCCCAGG
GGGATCTCCTAAGTGGCAAGGCCTGTTGGGAGGGGCTGGTGAGAGGCCACTCTGGCGGGA
AGACCCCCAGCCACCTGGAGCCCCTAGCCACTGCCTGCTGCGGCTCCCTAGGGATCCAGG
GCCATCAGAGAAGCTCCAGCGACACTGTTTATTTTCAAATGACACTTTTTAAGAAAAACA
GCCTCACCCAAATGCTTGGCCCTGAGTCTGGAATGTGCAGACAGACAGCTGCCCCTCCCC
AGAGCCTGCACGGCCCTCCGGGTGGGGAGGAGCAGGGGGCACCCCTGGGACCGGGCCGC
AGGCTGTCAGGGCACGGAACGTGTCTCTGGGCCCTGTCCTCAATTCCCGGTGCCCAGTGG
CCCCAACTTCCCAGCAGACCCAGCAGGGCCCCAGCTTGTCTTGGCCTGGCCGCTGGTCCT
GTCACCCCAGGCCTGGAGTTCTGGAAGATTCTGCTCCTGCTCCCGTGTGCACATACCACT
CCCCGGGGCAGCCCTGCACTTCTGTTCCTGCTGGGCTCCCTGCCTGCATCCGTGAGGCCT
GCAGCCCGCCTGATCTTCCAGGTCCTCCTCCGAGCCCCCGCCTCCAGGAAGCCCTCCAGG
AGAGCTCAGGAGGGTCGGCTCCCTGCGCGCAGCTGTCAGACCCCTGGGCCCACCCCGCCG
GCTGCTAGGGTCCAGGTTCCCCACAAGCCCTCGGGCAGAGGCTGGGCCGCTGGTCCCTC
GGAGACAACTGGCTCCGAGGCCTTGCCCTAGACGGGTTTCCGGGAGCCCGTCCCCAGCGG

FIGURE 8, CONTD.

CACCCACTGAGTTTTGAACACTTGGCGCCACCCCCACACCCCAGGCGGTGGCCAGGAGGC
CTCCTGGGCAGCAGACAGTCCGTGAGGTGGCCCTGGGGTGGCTCCTGACCTGGGCGCTGG
CCCAGCCCTGGGCACAGCTTTCCAGATCTTGCCTGCCGCTTCCTCCAGGCTGCCTCGGCC
CCTCCCGCCTGGGGGTGCCCAGCTTTTCCTGGAGGATGCCCACCCTTGCCCATGGTCAGG
GAGGGGCTGAGAAACCCCACCTCGTGCCTCTGCCCGGCCTATGCCAGGGGAACCAGGTTC
CCTCCCGCAGGAGGGGACCGAGTCCCTGACAGCCCACTGCAGAGGGGAGGAGGTGCCTGG
CTCTGCCCCCAGCCCCACCAACCCCGTGGCTTCCTGTTTCGCAGCCCACAAAGCACTAAA
GGCCGCAGGTCCTGGAACATCAAAGACCCGGGAAGTCCATTGTATTGAATTGAGTGTAAA
TGAGCCTGAGGCCTGTGGCTTGCGTTTCCCACAATTACCGCTGCCCGGGAAGGGCTCCGG
AACCGACACAGCCCCCAGGGCCCCTTGCCCATGTGGGGAGCCCAGGCTGGCCTGAAGAAG
CCCCATAAGGTGGACCCCACTTTGAGCCCCCACGAGAGTGGGCCAAGGACCAGGTCAGGG
GCTGCCCAGGCTCTGGGCCTCCTCTGCCTGCCAGGTGGGCTCCCTCGGGGCCCAGCCTGG
CCTGCAGGACCTTCCCACGCTGAGTTCCCCAGCCTGGTATGAGCGTAGTGGACGGCAGCC
ATGCCCAGCACTCAGGGGCCTGAGGGACAGAGCGGGAACTCCAGCCCCCGGGTCCTCGGC
CCCTAGGATCCTTCTAGGTGGGGAAGCCCAAGGGAGCAGAGGGGTGAACGCAGCTGTGTG
GGGCCCCAGGCTGCCGAGCAGACCCCTCCTGCTCCACTCCTCGGCCGAGTGGGCGCCGAG
ATGCCGGGGCAGTGCCATTTCCCAGGCGCCACCGGAGGCTCCCAGAGGGAGTGAGGCACG
AGCTGGGAGGGAGGGCGGGGGGGCTGGGGAGGCAGAGAGCGGAGGCCGGAGGCCGGTGAG
GAGGCCCGGAGGGGGGCCTGGAGTCAATGACCCAGGGATTATCGTGCTGGGTCTTTGCAAA
GTTGGCTGAGCAAACGCCGGAGCCAAGGGTCAGGGAGACGGGACTGGCGGGGCCCCGCGG
CCCCCTTTCCCCTTTCTGGAAAAAGCCTGTTTCCCAGGTCAAAATCCAGCTCATGATCCG
CCCCCTTTGGGACTGATGTTCAGAGGCCCAGTGGTCCCAGCACCTCTGTCCACCGCCCCC
CCCACGCTCCCGGGGCCGCCAACCCCTGTGGGCTGCGAGGTGCGGGCACCTCTCCCTTCG
AAGCAAAGCCCTGCCCTGCGTGGGCAGCGTGATTTCCTGCTTCTCTGGGGCTGCACTTTG
ACTGGGGTGGGGGGGTGG

Contig 4 (1592 bp)

AGCCCCTCAGCCCCTCCGAGCAGCTGCTGGGCTCAGCGGGCTCGCCCCCCGATGTGCGGC
CCTCCATAATCAATCATGGAGGGCCGGGCCCGGGGGGGCGGGCCGACCTGTCAGCCAGC
TCCAAGGGCAGGGACAGCTGCTGTTCCGGAGGGTTCCCAGGGGCCAGCCCCACCAGACAG
CGGCCTCGGCCCCCCTTCCCCGAGGGGCACCCCCACGGAGGGCCCAGACCGGAGGGACTC
GGGGCCCAGAGGCCAGGGCAAGAGTGAAGGCAGCGCCGGTGGGAGCGGCGGTCAGCGGGG
TCCAGGCTTCAGTTCCCAAGGAGCCCCATGCCCTGAGCCCGCACTGAGCCCTGTGCAGCC
TGTGGGTGCCGCCGAGGCCCGCCACCCCGCCCCCCACCAGCCTGGGGTCGAAGGAGGGAG
GGGGTGGCCTGACGGATGGTAACAGCTGCTCCCCCCACCTCGCCGGCGTGGACAGGGCTC
GCTTCTCCTGCCCGAGCCCCCGGCTGCCCCATCCGTCACGGCCCACCCAGGACTGTGCGT
CCAGCCTCCCTCCCTCCTAATCCCCCCGCATTTTCCGAATTCTCGGGCCACTGCTGCTTC
CTCCTCAAATTCCTGGCCCCCCTCGCCCCATCCCCGCCATGGGAAAGGGCCGCGATGCCA
GGACACTTGCTCGTCTCGGCCGGGCGGGGGGAGGAGCAGCTGGCTGGGCCCGGCAGCTGT
GAGGTGCGGGGGTGCCAGGGAGAAGGGCCCAGATTAGGGGGCGTCATGGGAAAGCTGGGA
GGGAACGCTACCCAGAGCCCCTCCTGCCGCAGCCTGTGCTGCTCCCTCTCCGCATTTCTG
GCCTCTGAGTGCTCCCTGGAGGGAAGGGACCACTGTGTCCTGCCGGCCTCTGGCTCTGCC
AGGAATGTCCATCTGTCCGGGCCGGGTTACCTGGCTCAGAGCGTGGGTACCAGCTCATCC
AGCCCTGACGCCTGCTCTCGGGAACAGTGGATGGGCCAGGCGCCCCCGTCACACCCCGCA
GCTGGGCTCCACAGACGGGCCCGGGATGGCCACGGAGGTGGGGGGCGGCCCCAGGGCGAG
GCTCCCTCCTGGAAGGGCTAGAGTGTGGGCTGCGCGGAGAGGGAGGCCGGACGGCCAGGC
CAGGTGCAGCCCGGGGCAGGTGCTGGTGGGGGCTGTGACCCACGTGTGCAGCTCAAGGGT
CCAGGAGCCCCAGGCACAGAGCCTCAGGGACAGACCCTCAGAGCCACAGCAGGAAGCCTG
GTGGCAGTAGCTGGCGGGGCCGTGGGGTGCTCGGCCCTGCAGACAGAGGCAGAGGCAGGC
TCCCTGCTGATGACAGGGGCTTTCTCTGTCCCCTGGGGGGCGGAGGGGGCCCGACCATGG
ACCCCGGGCCTCCTCTCGCACGATTCCCAGGCCAGCCTGGTCTCAGGCAGTCCAAGGTTG
CACAATGGTCTCCATCGTCCAGAGTTGCAGAGCCAGCACTCTCCCACTGGACGGCGGCCC
GGGGTGGGCTGCACCGCCGCTCAGGGCTCAGGGCCGCGGCCGGCCAGCCCNCCGCAGGCC
TTGACCCTGTCTCTTATACACATCTCAACCCTG

Contig 5 (831 bp)

TGAGATGTGTATAAGAGACAGGCCTTGACCCTGGGCCTGGCTCAGCTGCGCGCCCTCCTC
CTTGCAGCTCCGCCTCGACCCCATCCATCAGCCATTTTCCTACCCTTCCTGTAATAAAAA
ACCCGAAGCGGCGTGGCCCCGTGTCCGCTGGGGTGACTGCGGCCTGCCTGCTGGTGGCTC
CCACCTGGGCCCGGCCCCCTGAAAACACACACCCGGCGATGGCTTGCCCGGGGCCCTGGT
GGAGGGGCGGGGGGCCTCGCCTGCCTCTTGTCTGAAATTTTCGGTCCCACATGCCCCGAC
TCCTCTCCCGGCCCACCCTGCAGGCCCGGCCGGTGCCCCGGCCACTTTCCCGAAGGACGG

FIGURE 8, CONTD.

ACTCAGCATTTCCCAGGGCACCTGCTGATGGTGCCCAGACCCCGCGGGCCTTCCCGCCGG
GCGCGGCCCCACGTCGCCCCTCCAGTGGCCACAGCGGGCCTGGGCCAAGGCTGGGAGTTC
TGCACGGGCCTGGGGGAGGAAGGCGGGGGAGAGGGGACAGTCTCCTGGCGGGGACGAGGG
TGGGGGCAGCAGGTGGGGAGTTCCCACAGCCGGGGCAGCGGGACGCCGCTTGGCTGCCCT
GGGTCTCAGCCGGGGACAGTGCCCACCAGGAGAGAGACGGCAGACAGTACAGCCCACCCG
TTTTATATCCTCTCAGGCGGTCTGTGCTTTATTGGGGTAAATATGCAGGACATAGAAACT
CTGCCACTGGACCCCTTGGCCGGGGGACACAGCAGCGGCATTGCATGCTTTCTGGGTGCA
GCGCAGCCAGCACCACCGGCCAGAGCACCCCATCTTCCCGATCAACCGGAC

Contig 6 (4634 bp)

CTCTGGGCTAGCACCGTGGGGGCTTTGCCAGAGTGGAACTGAACTGGGTCCACCCCGGAG
CCCAGAGGGCGGTGAATGGGAGGCAGAGCCCATCCTGGGAATGGACCAGAAGAAAGGGAG
CGGGGGTGGGGGAAGGGGCATCAGATCCTGGTCCTTCCTTGTCGCCTGCGGTCCCTCTGC
CACCACTCCCCGAAGCTGATCTGGAGCACACGCGTCGTTAAAGCCGCCATCGAGGCCCCA
CTTCTGACAGACGGAAGGGGGCAGAGTGCCTTCCTCACCGGCCTCGCCCTGGGAAGGCCC
CTCCCTGCAGCCCAGGAAGCCAGCAGCAGGTGACAGAGCCAGGGGCCCAGGGCCCCAGGG
ACGGGCTCGCGCGCCCGAGCCGGGGGTCCCTTGGCGTCCCCATCCTCTCGTCCTGGAGCC
CTCCTGGGTGACCACAGGAATGTGCAAGGCGGCAGCCGGGTGGCGGCCGGGAGGCGGGTG
GGAGGCGGGCGGGGTGGCCTCTTCACGGGCGGGCCTGAGAGATGGGCGCCCGTCCGGCCC
TGGCGTCATCGTCTCCGCGTCTCTACCCACTGAGCAAAGACACACGAAATGAAGCTCGAA
CGAGCACAGCCAAAGAACGGCCGTTTCTGTCCTTTCTTCTTAATCCCTTTGGCTTAGGGT
TTCCCGGCCTGGACAGCCTGCCCAAGGGCACATGGGCATCCGTCCGGGGACATTCAGGCA
GTGACCAATCCCAGGCCACCCAGGCTGTGCCCTGCGTCGTGGGCCATTTCCCAGCCGGCC
AGAGATGGAGCAGCCACTGCGGGTCCCCGAGTCTCGGTGAGACAGTCAAGGATGGACCTT
GGATGGAGACCGGCGTGCGGCCATGTCCGTGGGTGAAGGAGGCGTGCAGGCCGTGCTGGG
GGACATGGTTGCTGTCCCCTCGGCCAAACCATGAAAAGCAGCCCTCTCCCCCAACCCCCA
GCACCAACCCGGAGACCACCCTCGGCCGGAGCCCAGCACGGCCACCGTCACGTCTCGGTC
GTCCAGCTTGGGACAGGTCAGTTCCCAGATGTCCAGGCTGGAGCTGGTCCTTGAAGATCC
TAGGGGTCCAGCCCAGCACAGGAGGGCCAGGTGAGAGCCCCCTGTGGTTCTAAGGATGCA
ACCAGGGGCCGGCGGGGTGCCTGCCCTAGAGGGGGTAACTCGGCCCCCTGGGGACCAGTC
ACCCCAGGAGGTCCCCAGAGCCCAGCTCGGAGGGCCACAGGTGCCCAGAGTCCCACCTGG
GGAAGGCTGCCCCTCCTGCCAGCCCCCGAGCCGGGCCCCTGGCGCCCGCGTCCAGCCGCG
ACCCCGGGGAGATATTCACCCCCTGCCCCCGTGAATCAGGAGGCCCCGAGCCCATGTTTT
CAGTCCTTTTCCTCCCATCCCAGCCCCCCAGGAGAAGAGGTGCTGAACTGGGTCCCCTGG
AGGCTCCTGAGCCCCAGAACAGTGCCCTCTGAGCAGACGGGCACTCTCAGACCAGCTCAC
GCTGGACAAGTCAGCTCCTGCCTGCCGCCTGATGGGCCCTTGGGAGAAGCAGACATGGTG
AGGAAAAGGCCCCTGTGCCCTTCACCCTAATTCCCCAGCCCCAAGTCCCACTGGGTTGCC
AGCTTCAACCTAAGCAAATAATTCGTGCCCTCTAAACAAACGCGCGGGAATCCCACCTGC
CCTTCCCCCGCCCGCCCCCC
ACCCCTGGCCTTGACCTCCAAAAGCACTTGAGGGGGCTTTCTCCAGACACCCTCCAACCC
CGACCCCATGAAGAAGGGGTGATGGGGCTGTTACCCCAACAAGCAAGAGAACGAAGCCCA
GAGAGGAGTTGGCGTGGACAGCAGGGGTCAGGCCCCTTTGCCCCGAGGGCAGGGCTGGTG
CCACCTGGGTCAGGCGGCAGGCCCTGGAAAAGCACCGGAAATGAGCACACCTGGGTCTCT
AGAAGGTTCTTCCAGACCTCTGGGGGCTGAGTCATTTCAACACTCCTGGGCCGGGCAGGG
CTTCTTCTTGGCCCCGAGGGACAAGGTCCCCTTCGTCCGGGGGGTACGGCCCCTGGACCC
CTGTCCCCCGCACCCCACCCTCCGCCTGGTGAGGGCCGCGGCCAGCTCTGGACACAGATC
CCTCAGAGCCCCTTCTCCCTCCCTGCTCCCTCGTCTTCCCAAGATGCCCCGGCCTCCAGG
TGGGGCAGCCAGGCGGCAGAATGTGGTCCAGGCCTCTCGGCCCCACCCACACCCCCCTGC
TCTGCCCTGACAGCCTCCAAGACGCAGGCACGTCGCTGCGTTCTGCGTCCTGTCTCCTCA
TGGCACAAAACGGTGCCCGCCTAGCTTCCCCCAGAGAAGGGAGATCGTGCTCCCCGGACG
GACCCTGCTCTGCCTGTCCTCCCGCCCGGCCTTCAGGGCCTCTCCCCAAGGGTGGCCGCG
AGGAGGCCCTCGCCTCCGGCCACGGGGGCTCCATCCTCCCGAGCCCGACAGGCCTCCGCC
TGGTGGTCCGACCTCTTCCCCAAGGCCCCGCCCATCCTCCTCGCGCTCCCCCAAACCCTG
CCTCTTTCCCCAGCGCCCTTGTCCCCACGGAAGACCCTCCACCCGTGCCATTACACGCTC
TCGCCCCACCCTCCCAGCCACCCCCCCTTCCCCATCCTCCTGGAAGCTCCCACTTCCTTC
CCGTCTCCCACGGCAGCAGAGGGTCAGCAGCTCAGGGGTCCTGGGGCCGTGGAGATGGCC
TGCCCGGGGGTCTCGCTGACCGCCTCCTACGGAAGCTGTGCCGGGGGGTGGGGGTGTCTC
TGCCCGAACGGCTGGAGGACGAGCCACATCCCAGGGCAGCCGGAACCTGCGTCCTGGTCT
GAGACGGAGAGGCTGGGTGCAGGTGGCTGAGGGGCCTGCACACAGCTTGGCCTGGGGTCC
CCTAGGTGACAACACTGGCTGAACACTCATTGCTGCTCCCCTTCCAGGGTGACCCTGGGG
TCCCCGTGTGGCCCTCAGGGCACACGGGGGCCCCACCAGGCCTCACAGAACCCCAGTGGG
ACTGCACCCAGGGCCCACAGAAGTGCGGGGGCACTGGGGGTCCAGAAACAACCCCACAAC

FIGURE 8, CONTD.

CAGGCCAAGGTGGCCAAGGCCTTACTCGAGCGGGGCTGCCCGTCCCAAGAGACTCTGGCC
AGTCGTCCGGATCCAGCTTCCCGGGGCCGGGCCGCCCGCTGGGCTCCAGGCGGTTCTGGG
GGGCCCTCCCCCGGGGTTCGCCCTCCGCTCTCAGCAGCAGGAAGAGGAGCGCGGCCAGC
GGATGGGGAGAAGAGGGCGCCCTGGCCATCTTGCTCCCCCTGGGACTTGAGGAGGGTCTC
GGGCCGGGCAGGCGGGACCGGGAGCCACAGAGACCCTGGAGGAGGCAGCATGGCGGGGAG
GTGACCGGGGAAGAGGGCCGTGTCCCAGGCTCACAGCCCGGCCTGGCCGCCCGGCCCTCG
GGAGGCGTGCCGCTGACCGCCTGGCCGGGAGGTTTGCTGCGTGTGGGGTTTGCAGAAAGT
GCTGAGCTGCTGAGCCCACAGGCCAGGCTCAGAGGGGACAGGAAGGAGGTTGCTGCCCAG
CCTCGGGCACTGCTGACCCATCTCCCGTTTCCAGGGCACCAGAGCCACCTAATCTGCCGG
CTCTGTGCCCAGGGACAGGCTTGCCTGATCTCTCAAGGCCGGGCGCTCCGCCTTCCCTGG
GAGAGGGTTAAACATCCAGCCCCAGCCAGCATCTCGGGCAGGTTCCTGGCTCCCCCCGCT
CGTGCCTCCTCTGAGACCCTGGTCGGCACACCTTTCCCTTGAGAGGAGGAGGAGGAGGAA
AGCGGATGGAACCAGTGACCCTGCAGCCCCTGAGGGCACCTTCCCACGTGCCCCCGCCCG
CCCCGCGTCCTCCGCCCCCAGTTCTCACGGCCCCAGTCCTGATGGAGGGAGGCGACCTC
CGGGCTCCCTGGCTCCCGCCGGCTCCGGAAGACAGGGCCGCTCGGCTGGGGCTGCAGGGA
GGGGCCCGAGACGCAGGAGAGCAGCCCGGAGGCAAACCCCGCGGCTCTTCCAGAAGGAGG
CCTGGCAGGGGGAGGGGGGTGCCACCACTGCTGTCCCTCTCGTGCCACAGTGGAGGGTGT
GGGTGGGCAGTGCCGGGGTGGGAAGTGCAGAAAGACCCTGGACCGTGGGGCTGGGCCGCC
ACGGGGGAGCGGGGTCTGTCAGGGACCCTGGGGGAGGGAGGCGAAGGGCTGGGGCAGAGG
CCGGATCACTTCCAGATTTGCTGTGGGACCAAGGGCCGGACCTCGGGGTGACTTCTTTTG
TGTGCTGGCCACAGGGGGGCCCCGGCGAGGTCACACGGAAGGGGGCTTCGGACCTGGCCT
AACAAGCCCACTCCCGAGGAAGATGCAAGGGGAGGCAGACGGAAGGGCCGAAGGGGGCGA
TCGGGGGACACCGCGGCAGGGCCGGGGCAGAGAAGGGAGGCAGAGGGCAGAGAAGGGAGG
CAGAGGGCACAGAAGGGAGGCAGAGGGGCCACATGCTTGGAGGGCCAGGGAGGAGCGGGA
ACGGCGTCCGGCGTCCAGCGCCGAATCAGGCCCGTCAGGCGGAGGGTGCGTGGACCTGCC
TGGCCTTCACGAGCACAGTCAGCAGGCTGTCTCTTATACACATCTCAACCATCAT

Contig 7 (482 bp)

AGCAATGGGGCCGTGACCTAAGGAGGCAGGCCCAGGTCAGTGGGGTGACCTCTCGTGGCC
CCGATGTTTGGAAATCCCCAAATCAAAATGACCCATCCGACAAGCTTGCATGCCTGCAGG
TCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCGCCCTATAGTGAGTCGTATTAC
AATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTT
AATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACC
GATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTT
CTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGC
TCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGAACCCC
TT

FIGURE 9

Human clone af087017.em_huml: H19 gene + flanking sequences

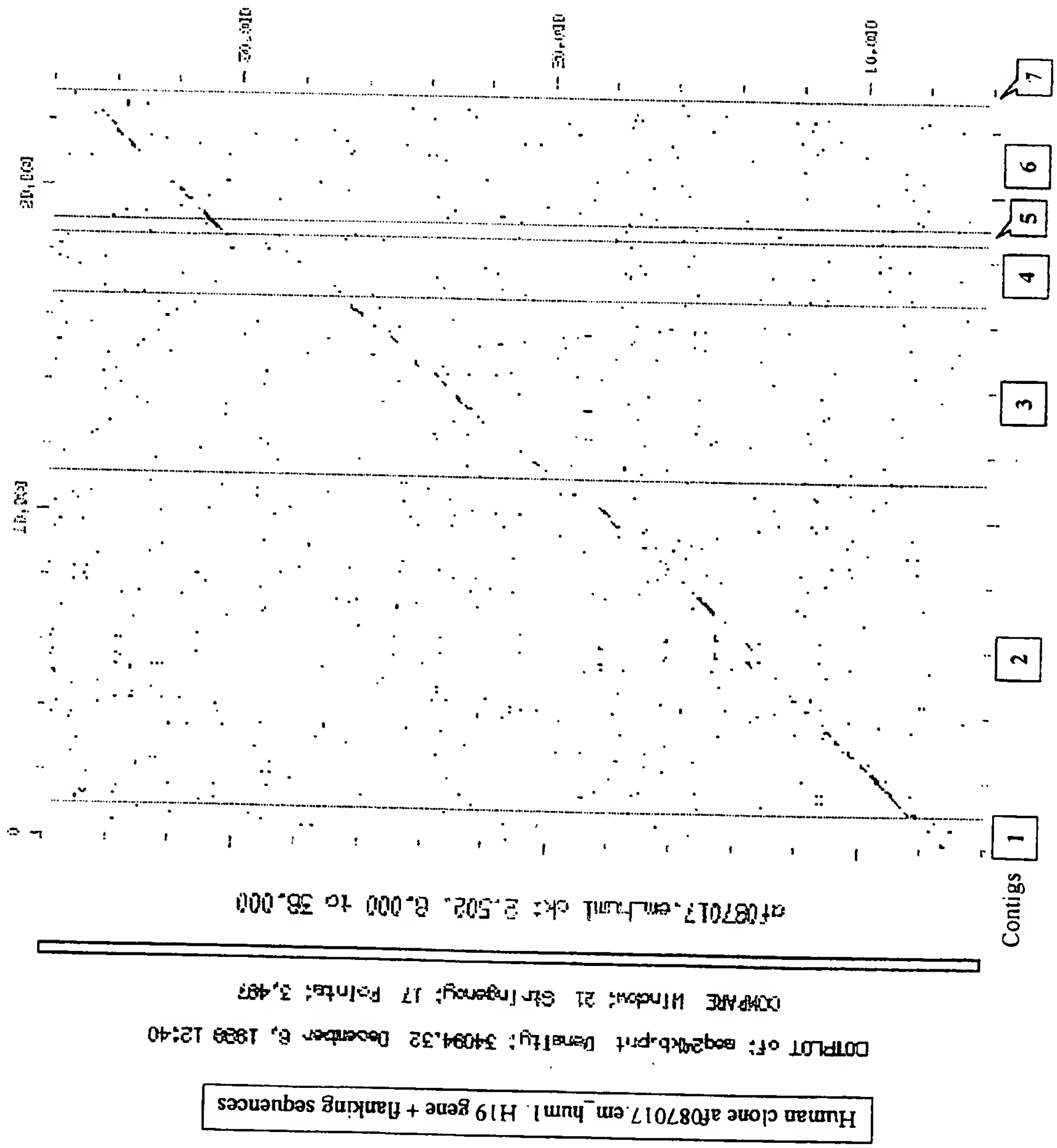

FIGURE 10

IDENTIFIED POLYMORPHISMS:

POLYMORPHISMS TYROSINE HYDROXYLASE GENE - CONTIG C3 (figure 6)

1	GGATCCAGCC(**A:T**)GCAGCC	1081 bp
2	ACAACCCCC(**-:C**)TCCCACAG	1149 bp
3	TGCGGAGGGG(**A:G**)GACCTG	1186 bp
4	AGGT(**CAAGGCCAGGT:-**)CGAGG	1210 bp

POLYMORPHISMS INSULIN-IGF2 - CONTIG C4 (figure 6)

5	CCC(**C:A**)CCCC(A:C)CGCCGC	438 bp
6	CCC(C:A)CCCC(**A:C**)CGCCGC	443 bp
7	CGCCGCAGCA(**G:A**)GCCG	455 bp
8	GCTTATGG(**G:A**)GCCGGG	503 bp
9	CACGGC(**T:C**)TC(G:A)GAGCA	525 bp
10	CACGGC(T:C)TC(**G:A**)GAGCA	528 bp
11	GTCTGC(**A:G**)GGCAGGTG	571 bp
12	CAAGCCCGG(**G:T**)CGGTT	636 bp
13	ACCTC(**A:G**)AGGCCCCCA	710 bp
14	GC(**C:T**)GGGCCCAGCCGC	867 bp
15	ACCAGCTG(**C:T**)GTTCCC	903 bp
16	GGC(**C:G**)CTCTGGGCGCC	1148 bp
17	GGGGG(**C:T**)GTCCCGGGA	1305 bp

FIGURE 10, CONTD.

18	GCGGT **(C:T)** GGGGGAGTT	1320 bp
19	CGCCC **(C:T)** GGTCCCGCT	1400 bp
20	TCCC **(G:A)** TCTGCCGGCC	1519 bp
21	GA **(T:A)** GCCCCATCCCCC	1547 bp
22	GG **(C:T)** GGCTGCTGCGGC	1607 bp
23	TGGCTGC **(G:A)** GTCTGGG	2222 bp

POLYMORPHISMES IN CODING REGION - CONTIG C10 (figure 6)

24	GCGCA **(G:T)** TGATTGGCA	341 bp
25	CGCCCCCCCCC **(-:C)** (G:C) GG	2247 bp
26	CGCCCCCCCCC (-:C) **(G:C)** GG	2248 bp
27	GCAGCCGGCTC **(C:T)** TGG	2257 bp
28	GTTGTTG **(C:T)** TCTGGGA	2413 bp

MICROSATELLITES

29 *PIGQTL1*: $(AT)^{11}$ 112 to 133 bp Contig 57

30 *PIGQTL2*: $(GT)^{8}$GCACGCGTGTGCGTGTGTAC$(GT)^{17}$ 1074 to 1144 bp Contig 95

31 *PIGQTL3*: $(CA)^{19}$ 223 to 260 bp Contig 105

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12Q 1/68, C07K 14/65, A01K 67/02 A3

(11) International Publication Number: WO 00/36143

(43) International Publication Date: 22 June 2000 (22.06.00)

(21) International Application Number: PCT/EP99/10209

(22) International Filing Date: 16 December 1999 (16.12.99)

(30) Priority Data:
98204291.3 16 December 1998 (16.12.98) EP

(71) Applicants (*for all designated States except US*): UNIVERSITY OF LIEGE [BE/BE]; 20 Bd de Colonster, B–4000 Liege (BE). MELICA HB [SE/SE]; Andersson, Leif, Bergagatan 30, S–752 39 Uppsala (SE). SEGHERSGENTEC N.V. [BE/BE]; Kapelbaan 15, B–9255 Buggenhout (BE).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): ANDERSSON, Leif [SE/SE]; Bergagatan 30, S–752 39 Uppsala (SE). GEORGES, Michel [BE/BE]; Rue Vieux Tige 24, B–3161 Villers-aux-Tours (BE). SPINCEMAILLE, Geert [BE/BE]; Sint Denijsstraat 26, B–8550 Zwevegem (BE).

(74) Agent: OTTEVANGERS, S., U.; Vereenigde, Nieuwe Parklaan 97, NL–2587 BN The Hague (NL).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report: 26 October 2000 (26.10.00)

(54) Title: SELECTING ANIMALS FOR PARENTALLY IMPRINTED TRAITS

(57) Abstract

The invention relates to methods to select breeding animals or animals destined for slaughter for having desired genotypic or potential phenotypic properties, in particular related to muscle mass and/or fat deposition. The invention provides a method for selecting a pig for having desired genotypic or potential phenotypic properties comprising testing a sample from said pig for the presence of a quantitative trait locus (QTL) located at a Sus scrofa chromosome 2 mapping at position 2p1.7.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 99/10209

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12Q1/68 C07K14/65 A01K67/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, MEDLINE, CHEM ABS Data, EMBASE, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ANDERSSON-EKLUND ET AL.: "MAPPING QUANTITATIVE LOCI FOR CARCASS AND MEAT QUALITY TRAITS IN A WILD BOAR x LARGE WHITE INTERCROSS" J.ANIM.SCI., vol. 76, 1998, pages 694-700, XP002104406 cited in the application	1-3, 10-12
Y	See page 696, "Carcass Composition" and page 698, Fig. 1b. the whole document	4-9, 13-27

	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
1 August 2000	08/08/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hagenmaier, S

1

INTERNATIONAL SEARCH REPORT

Int. ...tional Application No
PCT/EP 99/10209

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category°	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	KOVACS AND KLÖTING: "MAPPING OF QUANTITATIVE TRAIT LOCI FOR BODY WEIGHT ON CHROMOSOMS 1 AND 4 IN THE RAT" BIOCHEMISTRY AND MOLECULAR BIOLOGY INTERNATIONAL, vol. 44, no. 2, February 1998 (1998-02), pages 399-405, XP002104407	1,2,10, 11
Y	the whole document	4-9, 13-27

Y	JOHANSSON ET AL.: "COMPARATIVE MAPPING REVEALS EXTENSIVE LINKAGE CONSERVATION-BUT WITH GENE ORDER REARRAGEMENTS-BETWEEN THE PIG AND THE HUMAN GENOMES" GENOMICS, vol. 25, 1995, pages 682-690, XP000610181 See Fig.1, pig chromosome 2 the whole document	4-9, 13-27

Y	REIK W ET AL: "IMPRINTING IN CLUSTERS: LESSONS FROM BECKWITH-WIEDEMANN SYNDROME" TRENDS IN GENETICS, vol. 13, no. 8, 1 August 1997 (1997-08-01), page 330-334 XP004084608 Igf2 the whole document	4-9, 13-27

Y	CATCHPOLE AND ENGSTRÖM: "NUCLEOTIDE SEQUENCE OF A PORCINE INSULINE-LIKE GROWTH FACTOR II cDNA" NUCLEIC ACIDS RESEARCH, vol. 18, no. 21, 1990, page 6430 XP002104409 cited in the application the whole document	15

A	ANDERSSON L ET AL: "GENETIC MAPPING OF QUANTITATIVE TRAIT LOCI FOR GROWTH AND FATNESS IN PIGS" SCIENCE, vol. 263, 25 March 1994 (1994-03-25), pages 1771-1774, XP002018359 cited in the application the whole document	

A	KNOTT ET AL.: "MULTIPLE MARKER MAPPING OF QUANTITATIVE TRAIT LOCI IN A CROSS BETWEEN OUTBRED WILD BOAR AND LARGE WHITE PIGS" GENETICS, vol. 149, June 1998 (1998-06), pages 1069-1080, XP002104410 cited in the application the whole document	

	-/--	

1

INTERNATIONAL SEARCH REPORT

Int. ...tional Application No
PCT/EP 99/10209

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication where appropriate, of the relevant passages	Relevant to claim No.
A	WO 98 03682 A (UNIV IOWA RES FOUND) 29 January 1998 (1998-01-29) the whole document ---	
P,X	JEON ET AL.: "A PATERNALLY EXPRESSED QTL AFFECTING SKELETAL AND CARDIAC MUSCLE MASS IN PIGS MAPS TO THE IGF2 LOCUS" NAT.GENET., vol. 21, February 1999 (1999-02), pages 157-158, XP002104411 the whole document ---	1-27
P,X	NEZER ET AL.: "AN IMPRINTED QTL WITH MAJOR EFFECT ON MUSCLE MASS AND FAT DEPOSITION MAPS TO THE IGF2 LOCUS IN PIGS" NAT.GENET., vol. 21, February 1999 (1999-02), pages 155-156, XP002104412 the whole document -----	1-27

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte   tional Application No
PCT/EP 99/10209

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9803682 A	29-01-1998	US 5935784 A	10-08-1999
		AU 3513297 A	10-02-1998
		BR 9710875 A	11-01-2000
		CN 1230227 A	29-09-1999
		CZ 9900161 A	16-06-1999
		EP 0958376 A	24-11-1999
		PL 331353 A	05-07-1999
		US 5939264 A	17-08-1999